

Gly Gly Arg Arg Xaa Phe
20

(2) INFORMATION FOR SEQ ID NO:130:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Asp Pro Gly Tyr Thr Pro Gly
1 5

(2) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(iv) FEATURES:

(D) OTHER INFORMATION: /note= "The Second Residue Can Be Either a
Pro or Thr"

(v) SEQUENCE DESCRIPTION: SEQ ID NO:131:

Xaa Xaa Gly Phe Thr Gly Pro Gin Phe Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(iv) FEATURES:

(D) OTHER INFORMATION: /note= "The Third Positive Can Be Either a
Gln or Lys"

(v) SEQUENCE DESCRIPTION: SEQ ID NO:132:

Xaa Pro Xaa Val Thr Ala Tyr Ala Gly
1 5

(2) INFORMATION FOR SEQ ID NO:133:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Xaa Xaa Xaa Gln Lys Pro Phe Leu Arg
1 5

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

Xaa Asp Ser Gln Lys Ser Ala Thr Ile Lys Val Phe Asp Ala Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asn Leu Thr Ala Asp
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(86) SCIENCE INSPIRATION: SEQ 10 880:135:

Ala Pro Glu Ser Gly Ala Gly Leu Gly Ciy Thr Val Gin Ala Gly
1 2 3 4 5 6 7 8 9 10 11 12

420 INFORMATION FOR 600 TO 800,000

14) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acids
(C) STRANDEDNESS:
(D) TOPOLACY: linear

4833 SEQUENCE DESCRIPTION: SEQ ID NO: 137:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile
1 5 10 15

Ann. Vol. I. Ris. Lenz. Vol.
70

(2) INFORMATION FOR SSO ID NO: 128

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

533) MOLECULAR TYPE: DNA (genomic)

Table 1: Sequence description: seq ID: 10-125.

GCGACGCTGT CCTTGAGTTT CGGTTGATGG GTTTCGCGTC OCTTGACGTC GCGGCGGGCGG
 60
 TCACCATGCG ACCGACGCGG GCGTCAAAC CGTTGAGGG ACACCAAAAC GCGGAGCCCG
 120
 GGAAGTTCT AT CGCGTTGTTG CGCGCGACG AGCGAGCGC CGTCGCGGG CCTTGCGCGG
 180
 ATGATGCCAC CGTGGATTG CAGGCGCGA CGATTCGGG TTGAGAAC ATGATGCCAC
 240
 GCGTGGTAC CTGAGCGCG CGCTGCTGGG CGCGCGCTTC CGCTGCGGG CGACGGCGGG
 300
 CGGTGCGCGG TTTTGCGT CGCGCGCGT CGATCGCGT CGCGCGCGT ATTCGCGCGT
 360
 TCGCGCGCGT CGACCGCTGA ATCGCGCGA TCGCGCGCG CGCGCGCG CGCGCGCG
 420
 CGCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCG CGCGCGCG CGCGCGCG
 480
 CGACCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCG CGCGCGCG CGCGCGCG
 540
 CGACCGCGCG CGACCGCGCG CGCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCG
 600
 CGCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCG CGCGCGCG CGCGCGCG
 660

GGACGCCAGA GCGTACGGAA GAACCCACCC AACGATGTCG GAAACCAAGTG CAGACCCGTGG	726
CCCCCAGAAC GCTYGGGGCGG GCTGGCGACG CGCGGCGGGG TGCGGCGAAC GCGAGCGCG	780
GGCGGCACCTT ATTCGGCGCG TTTCGATGAC GCTGGCGGCTT TCACTAGCGTT CGGAGCGCAT	840
GGCGGGTGT GCGGTTGAGCG TGTGACTGCC CTGTCCTCAC GA	892

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 815 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:139:

GAATTAACCA ATGGCTCGCG CGCGCCCGGC CGCGCGATCC GCGCTGGCGG CGACCGCCCG	80
CGATGCTTAC GGTGCCCCCG TGGCGCTTCT CGCGCGCGTC CGCGCGAGCG GCGTGGGTC	120
CTAGGCGCT GTTACCGCCC TGGTTGGCGA CGACCGCCCG CGACCGCCCG GTAGCGCGCA	160
TGGCGCGCTT CGCGCGCGCG CGACCGCTGG CGCGCGCGCG AGCGTGGCGA CGCTTGGCGA	240
CGACCGCCCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	300
TGGTGGCGCT AGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	360
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	420
GGAGTGGCG CGATTAGGCGA CTGACCGCCC CGACCGCGCG AGCGTGGCGT GGTGCGCGCG	480
CACTTCCGA CGACACACCA GGAGGGGTTT TGCGGGGAG TGCGTGAAT CGCGCGCGAT	540
AGCGCGTACG TGTGGCGCTG CGTCGACCGT GCGTATGATG TCGAGCGTAA CGTGACCGCG	600
CGCGCGCGCG CGCGCGCGCG CGACCGCGCG AGCGTGGCGT GGTGCGCGCG CGCGCGCGCG	660
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	720
AGCGTGGCG CGGTGGCGCTG TTTCGCGCTT CGTCGACCGT AGCGTGGCGA CGTGTGCGCC	780
GGCGAAGGCA CGACACACCA CGTGGCGTCA AGCGCGCGCG AGCGCGCGCG CGGTGGCGCG	840

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1152 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

ACCGGGGCGC	GGCTGAGTTC	TGAGATCTGA	GACTCTUGG	ACTTCAGGG	GGGTTTACCC	60
CTTCTCCCA	GGACAGCTCT	GGAGATCTC	CCCCGGAA	CAAGCCCTCA	TTTGAGCGTC	120
TATGACCGT	TGAGACGAA	GGTCATCCG	GGAGTTGATA	TGAGACCGT	GGGCTAACAG	180
GTGGGCGA	TTGTCAGCT	GTATGCTCG	GACTCCCT	GGGGGATCA	CTTTCGCCAC	240
GGCCGGTGA	TGTTTGTAG	CCAGAGCTC	GGGAGGCC	AGTATCGAT	CGAGACGCTG	300
GGACGGATCA	GGGTTGTTGG	GGGGGCGAG	ATGACACGG	CTTTCATCGT	TGAGATGCTC	360
AGACGAGCG	GGCAGCTCA	GGCTTCA	ACCGGCGGC	ACTACCGAG	GGGGATCTCA	420
ACACCGAGC	TTCGATACCG	GGGGCGCTC	GGTCAGCAAG	TTTCAGGAC	GGGGATCTCT	480
GGTTTTCAC	TGTTGTTAG	CAAGCGCTC	GTCTGAGGA	AATATTTGAA	TGAGGAGGCG	540
TTGATTGGG	CACACAGCT	GGGGCAAGC	CTTGGTAGAG	GGATCCGCG	GATGAGGAC	600
TGCTGGCGT	GGGGGAGATG	ATGGGGCTC	CTTGGGAGCT	TGAGGGTT	GGGGGAAAT	660
GGGGGAAAG	CATACTTCAC	GGGGCTGGGG	CATTCCTCC	GGGAGGTT	GGGATTGCG	720
GGGGGCTCGA	CTGGGGCGCC	GGGGGGGG	TTCTACTGA	GGGGGAGCT	GGGGTATTG	780
CTGGGTTACA	AGAACATCT	AGGGGGGAGC	GGGGGCTCTA	GGGGGAGG	GTATATGGT	840
TTCTCTTAC	AGGATTCA	AGGGGGCGCA	AAGTCTGTC	GGGATTTGG	GGGGGGCTCC	900
GGTCAAACCG	CTGGGGGCT	GGGGGGCT	CTTACGTTAT	GGGGGGCTCC	GGGGGGGGCC	960
GGGGGGGGCA	ATACCGAGCG	ATGGGCGAG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1020
ATACCGGGGG	CACACTCCG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1080
GGGGTACTT	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1140
GGGGGGGG	AG					1152

(xii) INFORMATION FOR SEQ ID NO:148:

- (i) SEQUENCE CHARACTERISTICS:
- (ii) LENGTH: 655 base pairs
- (iii) TYPE: nucleic acid
- (iv) STRANDEDNESS: single
- (v) TOPOLOGY: linear

(iii) MOLECULE TYPE: DNA (genomic)

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTGGGGCGCA	TTGGGGAGGG	TCTACTTCG	GGGGGGTCTA	GGGGGGGG	TGCCCCACGAC	60
CACGATGCG	GGGGGGGGAC	GGGGGGGGCT	GGGGGGGG	GGGGGGGG	GGGGGGGG	120

CGGCTCGAAT CGGCCGTTGG CGCGCTTTCG CTGGCGCGG GCGATCGCG GTGTTTTCGCG 180
 CCTGCTTTGC CGCCGATTTG CGCGCGCGCG CGCGTGGCG CGCGGTATCG CGAGAATCGC 240
 ACGACACAC CGCGAGATCGC GGTCGTCGCA AGCTTTCGTA CGCTCGCGCG GGGCGCTTC 300
 CGCGGAAATT CTACTTGCGA GAACTCTTCG CGCGATCGCG CGCGCGCGA CGCGCGCGCG 360
 TCGAGCGCAC CGTCGATTCG CGATGGCGCG CGCGATCGCG CGCGACGCG CGTCGCGCGC 420
 TTGCGGACCG CGACCGCGTA GGTCGTCGAG TCGCGCTAC CGCTGGCGT TTGGCGGACCG 480
 TCGCGACCGT GTGTCGCGCGT CGCGCGCGAA AGCGCGCGCGT CGCGCGCGAT CGCGATCGCG 540
 TCGCGCGCGCG CGCGACTCGC CGCGACTCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 600
 TCGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 660

100 INFORMATION FOR GPO ID NO: 2021-00000000

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 267 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

LNU3 SEQUENCE DESCRIPTION: SPC 1D 402113

Asn	Ala	Val	Val	Ile	Phe	Ala	Val	Ile	Gly	Phe	Ala	Ser	Lys	Ala	Val
3				5				10						15	
Ala	Val	Ala	Val	Thr	Ile	Arg	Pro	Thr	Ala	Ala	Ser	Lys	Pro	Val	Glu
				20				25					30		
Gly	His	Gln	Asn	Ala	Gln	Pro	Gly	Lys	Phe	Met	Pro	Lys	Lys	Pro	Thr
					35			40					45		
Gln	Gln	Gln	Ala	Pro	Val	Pro	Pro	Phe	Pro	Pro	Asp	Asp	Pro	Thr	Ala
				50			55				60				
Gly	Phe	Gln	Gly	Gly	Thr	Ile	Pro	Ala	Val	Gln	Asn	Val	Val	Phe	Arg
	65				70				75					80	
Pro	Gly	Thr	Ser	Pro	Gly	Val	Gly	Gly	Thr	Pro	Ala	Ser	Pro	Ala	Pro
				85				90					95		
Gln	Ala	Pro	Ala	Val	Pro	Gly	Val	Val	Pro	Ala	Pro	Val	Pro	Ile	Pro
				100				105					110		
Val	Pro	Ile	Ile	Ile	Pro	Phe	Phe	Gly	Trp	Gln	Pro	Gly	Met	Pro	
				115				120					125		
Thr	Ile	Pro	Thr	Ala	Pro	Pro	Thr	Thr	Pro	Val	Thr	Thr	Ser	Ala	Thr
				130			135					140			
Thr	Pro	Pro	Thr	Thr	Pro	Pro	Ter	Ter	Pro	Val	Thr	Thr	Pro	Pro	Thr

150

146	150	155	160
Thr Pro Pro Thr Thr Pro Val Thr Thr Pro Pro Thr Thr Pro Pro Thr			
165		170	
Thr Pro Val Thr Thr Pro Pro Thr Thr Val Ala Pro Thr Thr Val Ala			
180		185	
			190
Pro Thr Thr Val Ala Pro Thr Thr Val Ala Pro Thr Thr Val Ala Pro			
195		200	
			205
Ala Thr Ala Thr Pro Thr Thr Val Ala Pro Gln Pro Thr Gln Gln Pro			
210		215	
			220
Thr Gln Gln Pro Thr Gln Gln Met Pro Thr Gln Gln Gln Thr Val Ala			
225		230	
			240
Pro Gln Thr Val Ala Pro Ala Pro Gln Pro Pro Ser Gly Gly Arg Asn			
245		250	
			255
Gly Ser Gly Gly Asp Leu Phe Gly Gly Phe			
260		265	

(2) INFORMATION FOR SEQ ID NO:143:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 174 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Ille Asn Gln Pro Leu Ala Pro Pro Ala Pro Pro Asp Pro Pro Ser Pro			
1	5	10	15
Pro Arg Pro Pro Val Pro Pro Val Pro Pro Pro Leu Pro Pro Ser Pro Pro			
20		25	
			30
Ser Pro Pro Thr Gly Trp Val Pro Arg Ala Leu Leu Pro Pro Trp Leu			
35		40	
			45
Ala Gly Thr Pro Pro Ala Pro Pro Val Pro Pro Met Ala Pro Leu Pro			
50		55	
			60
Pro Ala Ala Pro Leu Pro Pro Leu Pro Pro Leu Pro Pro Ieu Pro Thr			
65		70	
			75
Ser His Pro Pro Arg Pro Pro Pro Ala Pro Pro Ala Pro Pro Pro			
80		85	
			90
Ala Cys Pro Phe Val Pro Pro Val Pro Pro Ala Pro Pro Leu Pro Pro Ser			
100		105	
			110
Pro Pro Thr Gln Leu Pro Ala Asp Ala Ala Cys Pro Pro Ala Pro Pro			
115		120	
			125

Ala Pro Pro Leu Ala Pro Pro Ser Pro Pro Ala Gly Ser Ala Ala Ile
 130 135 140
 Arg Ala Lys Thr Gly Ala Thr Ser Ala Ser Thr Ileu Gly His Arg Ala
 145 150 155 160
 Leu Pro Asp Asp Thr Thr Ala Arg Gly Cys Arg Arg Thr Gly
 165 170

(2) INFORMATION FOR SEQ ID NO:144:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

Gin Pro Pro Ala Glu Val Ser Asp Gin Arg Val Ser Gly Ileu Thr Gly
 1 5 10 15
 Ala Val Gin Pro Ser Pro Arg Thr Thr Ala Glu Asp Pro Arg Pro Arg
 20 25 30
 Asn Arg Arg
 35

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 104 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

Arg Ala Asp Ser Ala Gly Cys Thr Cys Arg Trp Cys Asn Pro His Glu
 1 5 10 15
 Cys Arg Arg Pro Ala Met Arg Gin Gin His Gly Ser Arg Ser Thr Thr
 20 25 30
 Pro Pro Gly Pro Arg Gly Arg Ser Ala Arg Val Arg Pro Gly Arg Leu
 35 40 45
 Phe Pro Trp Ala Gly Ser Ser Asp Val Phe Pro Pro Trp Phe Ala Ala
 50 55 60
 Ile Met Pro Ala Arg Arg Val Gly Arg Pro Val Trp Pro Asn Val Asp
 65 70 75 80

Gin His Thr Arg Asp Thr Gly Leu Cys Lys Leu Phe Glu Arg Arg Ala
 89 90 91 92 93 94 95 96 97 98 99 100

Gly Gln Leu Arg Arg Gin Phe Tyr
 100

(2) INFORMATION FOR SEQ ID NO:146:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 55 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PCR primer"
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GGATCCATATGGGCGATCATGATTCATGCACGTGATGCA CATCATGGG ACC

53

(2) INFORMATION FOR SEQ ID NO:147:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 42 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PCR Primer"
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CCTGAATTCA GGCGCTGGTT GGCGCGCGCT CATCTTGAGC GA

42

(2) INFORMATION FOR SEQ ID NO:148:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PCR Primer"
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

GGATCTCCA GCTCCGAGC CACCGACCGG T

31

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTCTGAATTU AGGGUTGAA ATGCTGGGA T

31

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GGATCCAGCG CTGAGATGAA GACGGATGCC CCT

33

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GAGAGAAATTG TCAAGAAGCC ATTGGGAGG ACA

33

(D) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1983 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ORIGINAL SOURCE:

- (A) ORGANISM: Mycobacterium tuberculosis

(v) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 152..1273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

TGTTCTTCGA CGGCGGCCG GGGGAGGAGG 60000ACCGA ACAGCTTTC TCTTCGGCGA

60

ACCCATCCCGA AACCCCGCGA TAGCTGGCGG GACGTGCGG GUACGTCAAG GACGCCCGAC

126

AGCGGAAATTG AGAGACCGAG AAGGTTATGG C ATG ARA ATT CTT TTG CTT AGC

172

Val Lys Ile Arg Leu His Thr
1 2CTG TTG GCG CTG TTG ACC GCT GCG CGG CTC CTC CTA GCA GCG GCG GCG
Leu Lys Ile Val Ile Thr Ala Ala Pro Leu Leu Leu Ala Ala Ala Gly
18 15 20

220

TGT GGC TCG ARA CGA CGG AGC GGT TCG CCT GAA ACC GGC GGC GGC GGC
Cys Gly Ser Lys Pro Pro Ser Gly Ser Pro Glu Thr Gly Ala Gly Ala
25 30 35

266

GGT ACT GTC GCG ACT ACC CGG CGG TCG TCG CCG GTC AGC ATG TTG GCG GAG
Gly Thr Val Ala Thr Thr Pro Ala Ser Ser Pro Val Thr Leu Ala Glu
40 45 50 55

316

ACC GGT AGC AGC CTC CTC TAC CGG CTG TTC ATC CTG TGG GGT CGG GGC
Thr Gly Ser Thr Leu Ile Tyr Pro Leu Phe Asn Leu Tyr Gly Pro Ala
60 65 70

364

TTT CAC GAG AGG TAT CGG AAG GTC AGC ATC AGC GCT GAG GGC ACC GGT
Phe His Glu Arg Tyr Pro Leu Phe Asn Val Thr Ile Thr Ala Glu Gly Thr Gly
75 80 85

412

TCT GGT GCG CGG ATC GCG CGG TCG AAC GCG GTC AAC ATT GGG
Ser Gly Ala Gly Ile Ala Gin Ala Ala Ala Gly Thr Val Asn Ile Gly
90 95 100

460

GGC TCC GAC GCG TAT CTC TCG GAA GGT CAT ATG GCC GCG GAC DAG GGG
Ala Ser Asp Ala Tyr Ile Ser Glu Gly Asp Met Ala Ala His Lys Gly
105 110 115

508

CTG ATG AAG ATC CGG CTC GCG ATC TCC GGT CGG GAG GTC AAC TAC RAC
Leu Met Asn Ile Ala Leu Ala Ile Ser Ala Glu Glu Val Asn Tyr Asn
120 125 130 135

556

CTC CCC GGA GTG AGC GAG CSC CTC AAG CTG AAC GGA AAA GTC CTC CGG
 Leu Pro Gly Val Ser Glu His Leu Lys Leu Asp Glu Lys Val Leu Ala
 140 145 150
 604

GCC ATG TAC CGT GGC ACC ATC AAA ACC TGG GAC GAC CCG CAG ATC GCT
 Ala Met Tyr Glu Gly Thr Ile Lys Thr Trp Asp Asp Pro Glu Ile Ala
 155 160 165
 652

GCG CTC AAC CGC GUC GTG ANC CTG GCG GCG ACC GCG GTC GTT CGG CTC
 Ala Leu Asn Pro Glu Val Asn Leu Pro Glu Thr Ile Val Val Pro Leu
 170 175 180
 700

CAC GAC TCC GAC GGG TCC GGT GAC ACC TTC TTG TTC ACC CAG TAC CTC
 His Arg Ser Asp Gly Ser Gly Asp Thr Phe Leu Phe Thr Glu Tyr Leu
 185 190 195
 748

TCC AAG CAA CAT GLU GAG GGC TGG GGC RAG TGG GGC GGC TTC GGC ACC
 Ser Lys Glu Asp Pro Glu Gly Trp Gly Lys Ser Pro Glu Phe Gly Thr
 200 205 210 215
 786

ACC GTC GAC TTC GCG GCG GTG CGG GAT GCG CTG CGT GGT GAG AAC GGC AAC
 Thr Val Asp Phe Pro Ala Val Pro Glu Ala Leu Gly Glu Asn Gly Asn
 220 225 230
 814

GGC CGC ATG ATG ACC GGT TGC GCC GAG AAC CGG GGC TGC GTC GGC TAT
 Gly Gly Met Val Thr Gly Cys Ala Glu Thr Pro Glu Cys Val Ala Tyr
 235 240 245
 852

ATC GGC ATG AGU TTC CTC CAC CGG CCC ATG CAA CGG GCA CTC CTC GGC GAG
 Ile Gly Ile Ser Thr Leu Asp Glu Ala Ser Glu Arg Gly Leu Gly Glu
 250 255 260
 890

GGC CAA CTA GGC AAC ATG TCT GGC AAC ATT TTC TTG TTG CCC GAC GCG GAA
 Ala Gln Leu Gly Asn Ser Ser Glu Asn Phe Ile Leu Pro Asp Ala Glu
 265 270 275
 928

AGC ATT CGG CCT GCG GCG GCT GGC TTC GCG TCG AAC ACC CGG GCG AAC
 Ser Ile Glu Ala Ala His Asn Gly Phe Ala Ser Lys Thr Pro Ala Asn
 280 285 290 295
 1036

CAG GCG ATT TCG ATG ATC GAC GGC CCC GCG CGG GAC GGC TAC CGG ATC
 Glu Ala Ile Ser Met Ile Asp Glu Pro Ala Pro Asp Glu Tyr Pro Ile
 300 305 310
 1084

ATC AAC TAC GCG TAC GCC ATC CTC AAC AAC CGG CAA ATG GAC GAC CCC GCG
 Ile Asp Tyr Glu Tyr Ala Ile Val Asn Asn Arg Glu Lys Asp Ala Ala
 315 320 325
 1132

ACC GCG CGG ACC TTC CGG GCA TTT CTG CAC TGG GCG ATC ACC AGC GAC GGC
 Thr Ala Glu Thr Leu Glu Ala Phe Leu His Trp Ala Ile Thr Asp Glu
 330 335 340
 1180

AAC AAC GCG TCG TTC CTC GAC CGG GTT CAT TCG CAG CGG CTG CGG CGG
 Asn Lys Ala Ser Phe Leu Asp Glu Val His Phe Glu Pro Leu Pro Pro
 345 350 355
 1228

GCG ATG TTG AAC TTG TCT GAC GCG TTG ATC GCG ACC ATT TCC AGC
 Ala Val Val Lys Leu Ser Asp Ala Leu Ile Ala Thr Ile Ser Ser
 360 365 370
 1273

TAGCCCTGTT GACCCAGCAGG CGACAGCAGC CTGGCTGGCG CGATCGGGGT GCTTGGCGGA
 1393

GGTGTCTGGC CCGTCGCGGT GAGGTGAGGC CGCTTGGCC CGGATCGGG TOGTTGGGTG	1393
GTTATGCTGC GGTCAATCCCG CTGATTCGCG TGTGTTGGGT GCTGGTGCGT CTGGTGTACCG	1423
AGGCCATGGG TGCGATCGGG CTCAACGGGT TGCAATTCTT CGGGCGAACG AAATGGGATC	1513
CAGGAAACAC CTACGGCGAA ACGGTTGTCG CGACACGTCG CGGATCGGG TGCGCGCTA	1573
CTACGGGGG TTGGCGCTGA TGTGGGGAC GTGGGGACCT TGCGGATCG CGGGATGAT	1633
CGGGGTGCGG GTGTCGTGAG GGGGGCGCGT GGTTGTTGTCG GAGGGGTCG CGAAAGGCGT	1693
GGCGGAGGTT GTGGGAATGG TGTGGGATTT GCTGGGGCGA ATGGGGGGG TGGTGTCGG	1753
TTTGTTGGGGG GGTATTACGT TCGGGGCGTT GAGGGCTCAT GAGATGGTC CGGGTGATGG	1813
TCACACCGCT CGCGATGGCT CGGTGGTGA CTTACTGGGGG GGGGGCGCG GCAAGGGGGA	1873
GGCGGGTTG GTGTCGGTC TGCGGTTGGC GTGTGGTGGC GTTGGGATTA TGCGGACAC	1933
GACTCATGAC CGTGTGGGG AGGTGGCGGT GTTGGCGCGG GGGGGCGCG TGCGGATTC	1993

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:153:

Val Lys Ile Arg Leu His Thr Leu Leu Ala Val Leu Thr Ala Ala Pro			
1	5	10	15
Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser			
25	30	35	40
Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro His Ser			
35	40	45	
Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr Pro Leu			
50	55	60	
Phe Asn Leu Itp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr			
65	70	75	80
Ile Thr Ala Glu Itp Gly Ser Gly Ala Gly Ile Ala Glu Ala Ala			
85	90	95	
Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Glu Gly			
100	105	110	
Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser			
115	120	125	
Ala Glu Glu Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys			
130	135	140	

Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Glu Gly Thr Ile Lys Thr
 145 150 155 160
 Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro
 165 170 175
 Gly Thr Ala Val Val Pro Ile His Arg Ser Asp Gly Ser Gly Asp Thr
 180 185 190
 Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Glu Gly Trp Gly
 195 200 205
 Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ile Val Pro Gly
 210 215 220
 Ala Leu Gly Gln Asn Gly Asn Gly Gly Met Val Thr Gly Cys Ala Glu
 225 230 235 240
 Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala
 245 250 255
 Ser Gln Arg Gly Leu Gly Glu Ala Glu Leu Gly Asp Ser Ser Gly Asn
 260 265 270
 Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala Gly Phe
 275 280 285
 Ala Ser Ilys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro
 290 295 300
 Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn
 305 310 315 320
 Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu
 325 330 335
 Ala Thr Ala Ile Phe Asp Gly Asn Ilys Ser Phe Leu Asp Gln Val
 340 345 350
 His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu
 355 360 365
 Ile Ala Thr Ile Ser Ser
 370

(2) INFORMATION FOR SEQ ID NO:154:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1993 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:154:

TGTTCCTCGA CGGCAAGATG GTCGGAGGAA GACCCAGGCA ACAGCTGTTG TCTCTGGGGA

AGCATCCUGA AACCUCCGGA TACCTCGGGG GACTCTCGGG GGAGCTCAAG GACGCCAGC
 120
 CGGGAAATG AAGAGCAGC AAGGCTATGG CCGTAAATT CGTTTGCCATA CGCTGTGTC
 CGTTGTCAG CTGCTGCAGC TGTCTGTTAG AGCGGGGGG TGTTCTCTCA AACACCGGG
 180
 CGTTGGCCG GAAACGGGGG CGGGCGCCGG TACCTCGGGG ACTACCCCG CGTCTGCG
 240
 GGTGAGCTTG CGCGAGACCC GTAGZACCTG GCTCTACCGG CTGTTAACCT GTGAGGTTTC
 300
 GGGTTTCAC GAGAGTTCAT GAAAGTTCAC GATCACCGGT CGGGCACCCG GTTCTGTC
 360
 CGGGATCCGG CAGGCTCCGG CGGGACGCTT CAACATGGG CGCTCGAGGG CCTATCTTC
 420
 CGAAGGTTAT ATGGCGGGGG AGAAGGGCTT GATGANCATC CGCTTAACCA TTCTCGCTCA
 480
 CGAAGTCAAC TACACCCCTGC CGGGCTGGG CGAGCAGCTC AACGTCAGC GAAAGTCT
 540
 CGCGCTAT TACCGGGCA CGATCAAAAC CTGGGACGAC CGCGACGTCG CTGGGTCAT
 600
 CGCGGCGCTG AACCTGGCGG CGACGGGGT AGTTCGGCTG CACCGCTGG ACGGGTCGGG
 660
 TGACACCTTC TTGTTTTCAG AGTACCTGTC CAAGCAAT CCCGAGGGT GGCGCAACTC
 720
 CGCGGGCTTC CGAACACGGG TCGACTTCCC GGCGTGGGGG GGTGGCGGGG GTTACGACGG
 780
 CGAGGGGGCC ATGGTGAACG GTGGCGGGGA GACACCGGGC TGCGTGGCT ATATCGCGAT
 840
 CGACTTCCTC GAGAGGGCA CGTACGGGGG ACTCGGGAG CGCGAAGTAG CGTATAGCTC
 900
 TGACATTTC TTGTTGGGGG AGCGGCAAGG CTTCAGGGC CGGGGGGGT GTTGGGGTC
 960
 GAAACCGGG CGGAAACGGG CGATTTGAT GATCGAGGGG CGCGGGGGGG AGCGGTAAC
 1020
 GATCATCAC TACGTTTACG CGATGTCAT CGACCGGGAA AGGGGGGGG CGACCGGGCA
 1080
 GACCTTGAG CGATTTCTC AGTGGGGAT CGCGGACGGG AACGAGGGCTT CGTCTCGGG
 1140
 CGAGGTCAAT TTCCAGGGGG TGCCGGGGGG GGTGGTAAAC TTGCTTAAAG CGTGTGGGG
 1200
 GAGGGTTTC AGCTGGGGCTC CGTGGACCCAC AGCGGACAGC GACCTGGGGG CGGGCATGG
 1260
 GGTGCTTGC CGGATGATCTG GGCGGGGGGG GGTGGGGTGG CGCGGGGGGG CGGGGGGG
 1320
 CGGTGGTGG CGTACGGGGT CGCGGGGGGG CGCGGGGGGG CGGGGGGGGG CGGGGGGG
 1380
 GTCGCTGGCA TCGAGGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1440
 ACCGAATGCA ATCGGGGGCA CGCTTACGGG GAAGGGGGG TGACCGGGGG CGTGGGGGG
 1500
 CGGTGGGGGG CGTACGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1560
 TCGGGGGGGT CGTGGGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1620
 CGTGGGGGGT CGTGGGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1680
 TCGGGGGGGT CGTGGGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1740
 CGGTGGGGGGT CGGTTTCTGG GGGGGGATG GGTGGGGGGT CGTGGGGGGT CGTCAGGGCC
 1800
 CGCGGGGGGGT CGTGGGGGGT CGTGGGGGGT AGGCTTAACG CGTGGGGGGT CGTCAGGGCC
 1860
 CGGGCAACCG CGAGGGGGGGT CGTGGGGGGT CGTGGGGGGT CGTGGGGGGT CGTCAGGGCC
 1920

TTATCGCCAC CACCANTCAT GACGCTTTCG GGGAGGTGCC CGCTTGTCCC CGGGAAAGCCG	1980
CGATCGGGAA TTC	1993

(2) INFORMATION FOR SEQ ID NO:155:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 374 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

Met Lys Ile Arg Lys His Thr Leu Leu Ala Val Leu Thr Ala Asn Pro																																																																													
1	5	5	10	15	Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser		20	25	30	Gly Glu Thr Gly Ala Gly Ala Gly Thr Val Lys Thr Thr Pro Ala Ser		35	40	45	Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr Pro Leu		50	55	60	Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr		65	70	75	80	Ile Thr Ala Gin Gly Thr Gly Ser Gly Ala Gly Ile Ala Gin Ala Ala		85	90	95	Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Glu Gly		100	105	110	Asp Met Ala Asn His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser		115	120	125	Ala Gin Gin Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys		130	135	140	Leu Asn Gly Lys Val Leu Ala His Met Tyr Gin Gly Thr Ile Lys Thr		145	150	155	160	Trp Asp Asp Pro Gin Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro		165	170	175	Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly Asp Thr		180	185	190	Phe Leu Phe Thr Glu Tyr Leu Ser Lys Glu Asp Pro Gly Glu Gly Trp Gly		195	200	205	Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val Pro Gly		210	215	220	Ala Leu Gly Glu Asn Gly Asn Gly Met Val Thr Gly Lys Ala Glu		225	230	235	240
5	10	15																																																																											
Leu Leu Leu Ala Ala Ala Gly Cys Gly Ser Lys Pro Pro Ser Gly Ser																																																																													
20	25	30																																																																											
Gly Glu Thr Gly Ala Gly Ala Gly Thr Val Lys Thr Thr Pro Ala Ser																																																																													
35	40	45																																																																											
Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr Pro Leu																																																																													
50	55	60																																																																											
Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn Val Thr																																																																													
65	70	75	80																																																																										
Ile Thr Ala Gin Gly Thr Gly Ser Gly Ala Gly Ile Ala Gin Ala Ala																																																																													
85	90	95																																																																											
Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser Glu Gly																																																																													
100	105	110																																																																											
Asp Met Ala Asn His Lys Gly Leu Met Asn Ile Ala Leu Ala Ile Ser																																																																													
115	120	125																																																																											
Ala Gin Gin Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His Leu Lys																																																																													
130	135	140																																																																											
Leu Asn Gly Lys Val Leu Ala His Met Tyr Gin Gly Thr Ile Lys Thr																																																																													
145	150	155	160																																																																										
Trp Asp Asp Pro Gin Ile Ala Ala Leu Asn Pro Gly Val Asn Leu Pro																																																																													
165	170	175																																																																											
Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly Asp Thr																																																																													
180	185	190																																																																											
Phe Leu Phe Thr Glu Tyr Leu Ser Lys Glu Asp Pro Gly Glu Gly Trp Gly																																																																													
195	200	205																																																																											
Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val Pro Gly																																																																													
210	215	220																																																																											
Ala Leu Gly Glu Asn Gly Asn Gly Met Val Thr Gly Lys Ala Glu																																																																													
225	230	235	240																																																																										

Thr Pro Gly Cys Val Ile Tyr Ile Gly Ile Ser Phe Leu Asp Gln Ala
 245 250 255
 Ser Gin Arg Gly Leu Gly Gln Ile Gin Ile Gly Asn Ser Ser Gly Asn
 260 265 270
 Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala Gly Phe
 275 280 285
 Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp Gly Pro
 290 295 300
 Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile Val Asn
 305 310 315 320
 Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala Phe Leu
 325 330 335
 His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp Gln Val
 340 345 350
 His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp Ala Leu
 355 360 365
 Ile Ala Thr Ile Ser Ser
 370

(ii) INFORMATION FOR SEQ ID NO:156:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1777 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

```

GCTCTTGACCCACCGTGTGGG TGTGGGAGTT GGTGGCCGAGA TTGGAACTTCCTT GGTGACTGCTG 60
GGTGGGGGGG GCGCAACATG AGCGGAAAGG ATCGGAACTGG CGCGGGTGGG CGTGGGAGCT 120
GTAAGGAAACG GCGCAACGGGG CGCGGGTGGG CACGTTGGTT AGGGGGGTTT TCCACACAC 180
CTCGTGGGAGG CTGATGCCGTT GGAATTGGTC CGGGGGGAGG CTGGGGAGCA CGCGGGTGGG 240
CTGGGGGGGG CGACGGGGGG TGCTGGGGGG SCATGGGGTT CGCGGGAGCA GGTGGGGGGG 300
CGCAACGGCAT ACCATTTACGAA GACACCGCTT TGTATACTTT GTCAACGGCTT CGCGGGGGGG 360
AGGGGGGGGG AGGGGGGGGG ATGGCAACATTC TGGGGGGGGCA GACGGGGGGG AGGGGGGGGG 420
ACGGGGGGGG CGGGGGGGGG CTACGGGGGG CGGGGGGGGG TGGGGGGGGGG CGGGGGGGGG 480
ATGGCGTATA TGGGGGGGGG CGGGGGGGGG CGGGGGGGGG TGGGGGGGGGG CGGGGGGGGG 540
GGCTCTTACT TGTGGGGGGAA GATCAACGGG GACGGGGGGGG GGTGGGGGGGG GGTGGGGGGGG 600

```

TGCTCATCTGA	TCGGCGATGGCG	CGCGTGGGAC	AAGCTCAGCG	CATGGGGACC	CGACCGGAT	660
AGCTATGGCG	AGCAACGAGA	CTTTTGTGTC	GGCGCTCTGG	ATGGCGCTGA	GTCGGCGAC	720
CGCGTGGTAC	TGTTCTCTCA	GAATGGGGGC	TGGCGCGTC	GGTGGAGCTG	GGCTAAACAG	780
CTCGCGACG	GAOTCCAGGG	GATCCGGCTTC	ATGGAAAGCGA	TCCGTCACCC	GATGAGCTTG	840
GGGGGACTGG	CGGGCGCGT	GGGGGGTGTG	TCGCAAGGTT	TCCGATCGCC	TCAAGGCCAG	900
CGTATGCGCT	TGAGGACAA	CATTTTTC	GAACGGGGTC	TCGCGGGGCG	GAATCTGGA	960
CAGCTCACCG	ACGAGAAC	GAACCGTAT	CGCGCGCGAT	TCTGAGCG	CGCGGAGAC	1020
CGTGTGCGCA	CGTTGCGTG	CGCAAGAAC	CTTCGAATCG	ACGGTGAAC	CGCGGAGTC	1080
GTGCGTGTG	TCAGGAGTA	CGCGAGCTG	CTCGAGGAAA	CGCACATCG	GAACCTTG	1140
ATCAACCGG	AGCGCGGCG	GATCATCAG	GGCGCGATCC	CTGACTATGT	CGGGCGTGG	1200
CCNNCCNGA	CGAAATACAC	AGTCCCCGGC	GGGATTTCG	TTCAAGGAGA	CAAGCATGGC	1260
GTGCGTATGT	GGGGGGCG	TGGCGAGAT	GGCGAGCTG	GGACCGCTG	CATTTCACG	1320
GACCAAGAT	CTGATTCGG	CGAAAGCGG	CGCGCTGTT	GTGAACTGAT	AAAGCTTCC	1380
CGCTGGCGCA	BAGATTCTCA	CGGAAAGAGA	CAACCGCG	AGGYGCTTC	TCGCGAAGA	1440
GGTGCACAA	TATACTTGTG	AGGACRAAGG	TCTTCTTATT	TGCCAGGAGA	ATTAGCTCG	1500
GGCTTCTAT	GGGCTTACTT	CTAGGAACG	GAACGGAA	GGCGATCG	ATTGGACCTA	1560
TGGAACCGGT	ATUATGRRAG	CITCGAASCA	TTGGAAACAG	GGGGCTCG	GGGGCGTGG	1620
ATCATCCCAC	AGGGCTTC	TCACAAACCG	CGATGTAAT	ACGGTAACT	ACGGCGCACG	1680
GGCGATCGG	AGQAGGTGT	GGCGCGCTG	ACGAGGAG	GTATGGCG	GGTCTTGT	1740
TACGTGCCG	TTCACGATTG	GGGGGGCG	CGTCGCT			1777

(2) INFORMATION FOR SEQ ID NO:157:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GAGATTGAT	CGTACCGGGC	TCTTACGCG	CTCGTGUUG	TGATATCCCA	TATCACCGAC	60
GGCCATGTC	TGCGTGTGGA	CTTTCGCCCC	ATGGGGAGAC	CTTCTTAAC	CGAGGGTTG	120
ATCGATTAAT	GGGGGGAGG	CTTCGGGAA	GGGGGGCGGG	ATGTTGTTGA	GGGGGGCGGC	180
GGCGCGCTG	GAAGGGACCG	CTTGATGTC	ACGGCGAGTC	CGGGCGCTTA	GGCGCGCTTT	240

GGCGGCTGTC	GTCCACAGTC	GTACTCCGTT	GACGACCGGG	GGGGGTCCT	GGGTGAAGAC	393
CGTGAA	GGCGGCGAC	GGCGGCGATT	CAGA			324

(2) INFORMATION FOR SEQ ID NO:15B:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1338 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15B:						
GGGGTACCGC	CGCGGTTTGC	TGCAACGGGA	CCTGTAACGAC	CTGAACCTCT	TGCGGTTGGCG	60
AAACGATGTC	GAACGCTGCT	TGCGGCKCT	GTGCTGCTTC	AGGGTGTGCG	GTGATGCTGT	120
CGATGACCGG	CGGGCGACCG	GGCTTACGAC	CCTGCGAAC	GTGCGGCC	CGCGGAGGGA	180
GCACGACTTC	CAAGCGGACT	GTGCGGCT	GTGCGGCGCT	GTGCGGCT	TGACGCTGCG	240
GGAACTTCAC	GTTCGGGAGC	GTGCGGCTT	GTGCGGCGAC	GTGCGGCG	AGGGCGCGCG	300
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	360
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	420
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	480
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	540
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	600
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	660
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	720
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	780
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	840
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	900
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	960
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1020
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1080
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1140
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1200
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1260
GTGCGGCGCG	GGCAACGAC	GTGCGGCG	GTGCGGCGCG	GTGCGGCG	GTGCGGCGCG	1320

GCGGCCACCC CTACACCC

1338

(2) INFORMATION FOR SEQ ID NO:159:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 321 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

GCGGCCACCC CGCGCGACCC GCGGGTACCG CGCGCACCC CGCTGAACCC	60
TGGGCCTTCGG CGCGAACCCSC CGACCTGGGT TCGCTTGCGG CAAAGGCGGT AAAGCGCGAA	120
TAGTTGGGGCC CGCGCTGACA CGCGGGGCGCG CGCGGACCC CGCGCGCGC CGCGAGGTCG	180
GCACGCGCGG TGCGCGCGCG CGCGGCGACG AGCGCGACAC CGACCGCAAT CGCGGGGTA	240
AGCGCGGGCA CGCGCGGATC CGCGGTACCG CGCGGGCGCG CGCGCGCGCC CGACCGGGCA	300
ACGGGGGGCA CGCGGGGCGA CGCGGGCGCG CGCGGGGGCA CGCGGGGGCA CGACCGGGCA	360
ACGGGGGGCA CGCGGGGCGA CGCGGGGGCA CGCGGGGGCA CGACCGGGCA CGACCGGGCA	420

(2) INFORMATION FOR SEQ ID NO:160:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 452 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

GAGACGCCCT CGCGCGACCA TCGATCGCT CGCGCGATAC TTTCGGGAA CGCGACCGCC	60
CGCGCGTCCG CGTGTATCATC AGCGCGGGT AGCGCGCGAC CGCGACCGAA TGGCGTCTCG	120
CGTTCGCTTC CGAAGCTTC ACCTCGCGCG AGACCGCGAC CGACCGCGCA ATCGACAGGG	180
CGTGTCCACCA TTTCGGGTTCA AGAGATCTTCG TCGATGACTCT CGCGCGCGGA CGCTACCGCT	240
ACCGCGACCT TTGGGTTCAGC CGCTCGCCCG CGTGGGCGCG GATCGACCGG TTTCGGCGC	300
GGAGCGATAC CGCTCGCGGG CGCGAACCGA CGATCGGGG TTTCGGCGCG CGCGCGCG	360
TGGCGCGCGA CGCGCGCTAC GACCGCGCTG CAATCGAGGG CGACCGAGGG TATCTCTCA	420
ATCGATGTTG CGCGCGCGTC ACCGACAAAC CGACCGACCT CGCGGGCGCG ACACCGCGCA	480
ACCGCGCGCG CG	492

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 336 amino acids
- (B) TYPE: amino acids
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

Phe Ala Gln Ile Leu Val Glu Gly Asp Ala Val Gln Ile Trp Arg Ala
 1 5 10 15

Asn Ala Ala Asp Gin Ala Asp Pro Leu Gln Pro Gly Ser Ala Arg Arg
 20 25 30

Gin Arg Ala Ser Arg Ser Pro Arg Arg Leu Ala Gly Pro Asn Ala Tyr
 35 40 45

His Tyr Ser Asn Asn Arg Ser Ile Leu Cys Gln Arg Trp Pro Leu Pro
 50 55 60

Ser Ala Ala Glu Asp Val Ile Cys His Leu Cys Pro His Arg Gln Glu
 65 70 75 80

Pro Gly Leu Met Thr Ala Phe Gly Val Gln Pro Tyr Gly Gln Pro Lys
 85 90 95

Tyr Leu Glu Ile Ala Gly Lys Arg Met Ala Tyr Ile Asp Gln Gly Lys
 100 105 110

Gly Asp Ala Ile Val Phe Gln His Gly Asn Pro Thr Ser Ser Tyr Leu
 115 120 125

Trp Arg Asn Ile Met Pro His Leu Glu Gly Leu Gly Arg Leu Val Ala
 130 135 140

Cys Asp Leu Ile Gly Met Gly Ala Ser Asp Lys Leu Ser Pro Ser Gly
 145 150 155 160

Pro Asp Arg Tyr Ser Tyr Gly Glu Gln Arg Asp Phe Leu Phe Ala Leu
 165 170 175

Trp Asp Ala Leu Asp Leu Gly Asp His Val Val Leu Val Leu His Asp
 180 185 190

Trp Gly Ser Ala Leu Gly Phe Asp Trp Ala Lys Gln His Arg Asp Arg
 195 200 205

Val Gln Gly Ile Ala Phe Met Glu Ala Ile Val Thr Pro Met Thr Trp
 210 215 220

Ala Asp Trp Pro Pro Ala Val Arg Gly Val Phe Gln Gly Phe Arg Ser
 225 230 235 240

Pro Gln Gly Glu Pro Met Ala Leu Glu His Asn Ile Phe Val Glu Arg
 245 250 255

Val Leu Phe Gly Ala Ile Leu Arg Gln Leu Ser Asp Glu Gln Met Asn
 260 265 270

His Tyr Arg Arg Pro Phe Val Asn Gly Gly Glu Asp Asp Arg Pro Thr	275	280	285
Leu Ser Trp Pro Arg Asn Leu Pro Ile Asp Gly Glu Pro Ala Glu Val	290	295	300
Val Ala Leu Val Asn Glu Tyr Arg Ser Trp Leu Glu Glu Thr Asp Met	305	310	315
Pro Lys Leu Phe Ile Asn Ala Glu Pro Gly Ala Ile Ile Thr Gly Arg	320	325	330
Ile Arg Asp Tyr Val Arg Ser Trp Pro Asn Gln Thr Glu Ile Thr Val	335	340	345
Pro Gly Val His Phe Val Gin Glu Asp Ser Asp Gly Val Val Ser Trp	355	360	365
Ala Gly Ala Arg Gln His Arg Arg Pro Gly Ser Ala Ile Ile Ser Arg	370	375	380
Asp Gin Glu Cys Asp Phe Arg Arg Arg Asp Arg Pro Ala Cys Glu Leu	385	390	395
Ile Arg Leu Pro Ala Pro Gly Arg Asp Ser Gin Gly Lys Gly His Gln	405	410	415
Ser Gin Pro Leu Pro Ser Ser Gln Arg Gly Arg Gin Ile Tyr Val Ala Gly	420	425	430
Gln Arg Ser Ser Tyr Leu Pro Ser Glu Leu Val Ala Ala Phe Leu Trp	435	440	445
Ala Gin Phe Glu Glu Ala Glu Arg Ile Thr Arg Ile Arg Leu Asp Leu	450	455	460
Trp Asp Arg Tyr His Glu Ser Phe Glu Ser Leu Glu Glu Arg Gly Leu	465	470	475
Leu Arg Arg Pro Ile Ile Pro Glu Gly Cys Ser His Asn Ala His Met	485	490	495
Tyr Tyr Val Leu Leu Ala Pro Ser Ala Asp Arg Glu Glu Val Leu Ala	500	505	510
Arg Leu Thr Ser Glu Gly Ile Gly Ala Val Phe His Tyr Val Pro Leu	515	520	525
His Asp Ser Pro Ala Gly Arg Arg	530	535	

(2) INFORMATION FOR SEQ ID NO:162:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 204 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(x) SIGHTING DESCRIPTION: GRO ID NO: 1521

Asn Gln Ser Ala Pro Arg Ser Phe Met Leu Pro Ser Ala Arg Pro Arg
 1 5 10 15
 Tyr Asp Ala Ile Ala Val Leu Ieu Asn Glu Met His Ala Gly His Cys
 20 25 30
 Asp Thr Gly Leu Val Gly Pro Ala Pro Asp Ile Val Thr Asp Ala Ala
 35 40 45
 Gly Asp Asp Arg Ala Gly Leu Gly Val Asp Gln Gin Phe Arg His Val
 50 55 60
 Gly Phe Leu Glu Pro Ala Pro Val Leu Val Asp Gln Arg Asp Asp Leu
 65 70 75 80
 Gly Gly Leu Thr Val Asp Trp Lys Val Ser Trp Pro Arg Gln Arg Gly
 85 90 95
 Ala Thr Val Leu Ala Ala Val His Gln Trp Pro Pro Ile Val Val His
 100 105 110
 Phe Leu Val Ala Glu Leu Ser Gin Asp Arg Pro Gly Gln His Pro Phe
 115 120 125
 Asp Lys Asp Val Val Ile Glu Arg His Trp Leu Ala Leu Arg Arg Ser
 130 135 140
 Glu Thr Leu Glu His Thr Pro His Gly Arg Arg Pro Val Arg Pro Arg
 145 150 155 160
 His Arg Gly Asp Asp Arg Phe His Gln Arg Asp Phe Leu His Ser Val
 165 170 175
 Ala Met Leu Val Ser Pro Val Glu Ala Glu Arg Arg Ala Pro Val Val
 180 185 190
 Glu His Gin Tyr His Val Val Ala Glu Val Glu Arg Ile Pro Gln Arg
 195 200 205
 Glu Gin Lys Val Ser Leu Leu Ala Ile Ala Ile Ala Val Gly Ser Arg
 210 215 220
 Trp Ala Gln Leu Val Arg Arg Ala His Pro Asp Gln Ile Ala Gly His
 225 230 235 240
 Gln Pro Ala Gln Pro Phe Gln Val Arg His Asp Val Ala Pro Gln Val
 245 250 255
 Arg Arg Arg Gly Val Ala Val Leu Lys Asp Asp Gly Val Thr Leu Ala
 260 265 270
 Phe Val Asp Ile Arg His Ala Leu Pro Gly Asp Phe
 275 280

(2) INFORMATION FOR SEQ ID NO:25:

SEQUENCE CHARACTERISTICS:

(8) LENGTH: 264 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

ATGAAACATGT GGTGTTGGT GAGTCGAGC GCTTTGGCC GATTGCCCG CTAATCTGC	60
GGCTTGACG CGATCCCCGG TTTCCTCGAT GGTTCGGCC AAGGCGCCG GGTAAAGGAA	120
ATCCGCGTCT CGGTGATCA CTCGGCGCTG ACCCGAAGAC CCGTGTGTCG CACGTCGAC	180
CCCGCGGAGA TGCGCGCGC GTTTCGGAGC CTCAGCGCCG TTGGCTTCA CTGGGCGCC	240
CGAGCGGTGC TTGACGGTGT GGCG	264

(2) INFORMATION FOR SEQ ID NO:164:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1171 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

TATGGCGGA CTTGAGCGTC GCGGTCGAGG CGAACCGTT CAAKACCGG CGGACGAGG	60
AAGCCGGTGC GAPGTTAAC CGCGAACAG TGGTGAGCA CGGGGCTCC CGGGGAAAC	120
AGTGTCACG CAGCTGTTG CGCGCGCTGT GCTCGATTCG GGTTCGGAA TTGGGTTATC	180
TCTGCGTCA TGTGGCGGTG GGGCGCTCA TTATCGACT GCTGGATTC GGGGAGCTCG	240
CGCTTGGACG CTCGTTGATG TGGGAGCTG TTGAAATCGG TTTCGTCGG CGCTGAGTGT	300
TGGGGTGTAT CTCGCGCGAG CGTCGGGAGC GCGGAGCTG GGGGTTCT CGCGGCGGAA	360
ACCGCTCTG GACCGCGCGG GCGAACCTCC CGAAAGGAGC GCAAGTCGG ACAGTCGTC	420
ATCCCCAGCG GGCAGCGAGT TGCCTCTGT GCGGATTCG CGACGAGCT CGCGAGCGAC	480
CGGGATTCG CAGCGAGCG TTACCGAGTA CGGAGCTCA CTTCGCGAT CGCGAGCGCA	540
ATCTCTCGC CGGCGATCTG CAGATCCCGC TGTCGCGTC AGCTGAGACG CGCGAGCTG	600
GCGAGCGGT ATCGGAGATT GAGCGGAGC CGCAATTCTT CAATGGCTGC CGCGTCGCGC	660
ACTATTTGCA CTTCGCGCG TGCGCGATG TGAGGAGCA TGCGAGCTG CGAGGAGCTG	720
CGCCAGCTG CGACGCGGT AGCTCCCGGC CGGAGCGGA GGTGGGGGG GTGATCTTGT	780
CGGGCGCGT CGTACCGCTT GATCCACCGC TTTCGCGTGC CGGGGGGGAG CGCGATCGC	840

TTCAGGACCT	CGGGGTATGC	CCACGCCAAG	CFFGGGGGCT	TGTTGAGGT	CAGGAACCTC	908
ACCATCCGCA	CGGGGACCAA	GGTGGGGCAC	CTGACCTAAG	TGGGGGACGC	CGCATCGGC	966
GACTACAGCA	ACGTGGGCG	CTCCAGCTTG	TCTCTCTACT	ACGGGGTGTAC	GGTCCTAACGG	1030
GGGGACACCG	TGGGTTCGCA	CTTACGCC	GGGGGGGACA	CCATTTGGT	GGGGGGACGA	1080
ACCATCGGG	ACGGGGGCGTA	TACCGGGGCC	GGGGGGTGG	TGGGGGAGGA	TGGGGGGCG	1140
GGGGGGTGG	CAAGTGGGCG	GGGGGGCGA	C			1171

(2) INFORMATION FOR SEQ ID NO:165:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 277 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:165:

GCAGGGCGG	CAAGGGGGGG	GGGGGGGTTG	ACGGGGCTCGA	GGGGGGCTCA	GGGGGGCGAG	68
ACGGGGCGCA	AGGGGGGACG	GGGGGGACCG	GGGGGGAGGC	GGGGGGGGGC	GGGGGGGGGT	120
TGACCCAAAG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	180
GGGGGGACCG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	227

(2) INFORMATION FOR SEQ ID NO:166:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 304 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

CTCTGGTACG	ATGGGGGGGC	AGGGGGTGTAC	GGGGGGGGCC	GGGTGTACCC	GGGGGGCGCA	46
GGGGGGCGCA	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	120
GAATCTCCAA	CTGGGGGGCG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	180
GGGGGGGGAC	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	240
CAACGGTAC	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	GGGGGGGGGG	306
GGGC						304

(2) INFORMATION FOR SEQ ID NO:167:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1439 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(2) SEQUENCE DESCRIPTION: SEQ ID NO:168:

GTCGGACGCT	GCCTGGCTCG	TATACCAAGG	AACACATCGA	CCAGGGCGTC	CTGGGTGAAAC	60
TCATCGACGT	ATTTAACACT	GGCGCGTCA	GGGGCGCCGG	CGACGACCGG	GGCGGGGATC	120
TGATGGTGGA	GGTCTCTGAA	TACTTCTCTG	GCATTTCGG	TGGGGGGGAA	GGGAGCGGG	180
CTGGGAGTT	CTTTACCCG	CCAGCGCTGG	TCAAATGAT	CCTGGGATCTG	CTGGGGCGGT	240
GGAGTGGGGG	GGTGTATGAC	CGGCTGCTGG	CTTCGGGAGG	CATTTTCTG	CAGACCGGAA	300
AGTTATCTA	CGAACAGGAC	GGCATTCGGA	AGGATGCTTC	GATCTATGG	CGGGAAAGCA	360
TTGAGGAGC	CTGGGGGATG	GGGAAATTA	ACCTGGGAT	CCACGGATC	GACACRAGG	420
GGCTGGGCC	CCGATGGGT	GATACTTTC	CCCGCGGCA	GCACCGGG	GGGGAGGG	480
ACTACGGTAT	GGGCAATCG	CGGTTAAC	TCAAGGAGTC	GGGCGAAC	GGGGAGAC	540
CGGGCTGGG	CTTGCGCTTT	GGGGCGGCA	ATTAACGCTA	CTTGCGCTTG	ATTGGCAC	600
TCCTATACAA	CTTGGCGCTG	GGGGCTGGG	GGGGCGGTGT	GATGGCCAC	GGGGTGGATG	660
CTTGGAACTC	CAACCGCAAG	GGGGATATTC	GGGGCGAAC	CGTGGCGCG	GATTTGGTTT	720
CCCTGATTTT	GGGTTTACCC	ACCACAGCT	TCDGACCCAC	GGGATCCCG	CTGTGGCTTT	780
GGTTTTTC	CAAAACAAAG	GGGGGGGTTA	ACCAAGGTC	TATCAGCTTG	GGGGGGAGG	840
TCCTGTTTCT	GGGGCGCTCT	GAATGGGGC	ACCTATGTA	CCGGGGCGAC	GGGGCGCTGA	900
CTTGGAGKAA	GGGGCGGCA	ATCGGGGATA	CTTGGCGATC	GGGGCGGAGG	ACGGGGCAAC	960
CGGGCTGGG	TCATGGGGC	GGTATATGGG	GGACTGGCT	CTACCGGGCG	GGGGGGCTGT	1020
GGGGGGGGGG	GGGGCGCGG	GGTGTGGGG	GGTGTGGGG	GGGGCGCTG	GGGGGGGGGG	1080
ACGGGGCGAA	GGGGCGAAC	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1140
GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1200
GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1260
GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1320
GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1380
TCACGGTCAC	GGGGGGGGAC	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGGGG	1439

(2) INFORMATION FOR SEQ ID NO:168:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 329 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

GCGGCGGGCGG	GGGCGGAGATT	TCTCTGGCT	TGATTTGTCCG	TGGGGATAAC	GGGGGTGATS	60
GTGGTTCACGG	CGGGATGGCG	GGGGCTGGCG	GGGGCTGGCG	GGGGGGGGCG	GGGGGGGGCG	120
TGATTCGGCT	GTCTGGCGGC	CAAGGGGGCG	GGGGGGCGCG	GGGGGGGGCG	GGGGGGGGCG	180
GTGTTGGCGG	TGAGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	240
GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	300
GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	329

(2) INFORMATION FOR SEQ ID NO:168:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 80 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

GCAACGGTGG	CAAGGGCGGC	ACCGGCCACCA	CCCTGGGGAT	GGGGGGGGAT	AAGCTGTGGTG	60
CCGGGGGGCT	GATCGGGCAC					66

(2) INFORMATION FOR SEQ ID NO:170:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 392 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

GGGCTGTGTC	GGACTGACAC	GGGCTGTGTC	GGGGGGGGTG	GGGGGGGGAT	ATCCAGGCTGA	60
AGGCTACTA	CTTACCGCTG	GGGGGGGGCG	GGGGGGGGCT	GGGGGGGGCG	GGGGGGGGCG	120
TCAAGGCTAT	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	180

CCCCCCCCCA CCTTGCGCGC CAAAGGGGCTT CAACGCTCCG AACGGCGCGG ATCGTGTGAG	240
CACACTGTTG CTGGGGCAAC GAGATTCCTG CAAATGTCGA GCGCGGTGCG AGCGCAAGGA	366
CGGGGTATAC CTATGTCAC CTATGTCAC TCTAACACACG GCGATAACGA TCCCGTGTG	369
CGCGCACAGC CGAACGACTGC AAGACCGTTA CA	392

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

ACCGGGGCCA CGGGGGCGCG CGGGGTCGCCG AGTGCGGGCG CGGGGGCGCG	60
GCTATCGCGG CTGGGGGGCGG CACCAACGCCG TCTGGTGGGGG CTGGGGGGACG CGGGGGACAA	120
GGCGGGGGCGG CGGGGGCGCGG CGGGGGGGCGG GCGGATAACG CGATGGCGAT CGGGGGCGCG	180
GGCGGGGGCGG CGGGGGCGCGG CGGGGGGGCGG GGAGCGCGCG CGGGGGGGCGG CGGGGGCGCG	240
ACCGGGGCCA CGGGGGGCCA CGGGGGCGCG CGGGGGGGCGG CGGGGGGGCGG CGGGGGCGCG	300
GGTACGGGTG CTGGGGGGCGG TCTGTGTGGT CGGGGGGGCGG CTGGGGGGCGG TGGGACACCG	360
OCTACCGGGTG CGGGGGGGTT CGGGGGGGCGG CGGGGGGGCGG AAGGGGGGCC CGGGGGGGCGG	420
AGCGTGTGG CGGGGGGCCA CGGGGGCGCG CGGGGGGGCGG CTGGGGGGCGG CGGGGGGGCGG	480
ACCGGGGGGTG CGGGGGGGCGG AACGGGGGCC CGGGGGGGCGG CTGGGGGGTT CGGGGGGGCGG	540

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 699 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

CGACGCTCG CGGGGGGGATA CGGGGGGGCGC CGGGGGGGCGC ATGGGGGGCGA CGGGGGGGCG	60
GGCGGGGGTG CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCG	120
CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCG	180
CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCGC CGGGGGGGCG	240

CATGGCGGAT GGCCTTGCCG CGGGAAACAGA AGGAGGATC CTTGACTTCA CGGGCUACCT	360
CGAGCGCTG TGCGGCGCAC CGTCACGCT CGCGCGATC CAGCTGGCG AACCGGGCGA	366
TCTTGTTGCCG GGGGGGGCGG CGCGCGCGAC CGCGCGCGAC GTGCTGAACA CGCTGGCGAC	420
GATCATCTCA ACCAACATACG CGGTCTCTCT CGCGACCGTG GACCTGCGCG TCGCTGGTC	480
ACCACCTGCG CGCTGTACAC CACCCGAACG TTGCTCACCGC AACCTGGTGC CGGCGAATCTG	540
ATCAGCGCG TCGCTTATCG CGTGCAGGCC AGCGTAGTTT TGGCGCGAT CGTATAAGCG	600
CGGGCGTGGAA TTGCTCACCGC TCTCTGGCGC CGCGCTGGAC AGCGCTTCAA ACATCCAGGG	660
CGCTGGTCAAC TAAAGGATTG CGCGACGGCGAT	696

(2) INFORMATION FOR SEQ ID NO:173:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 497 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:173:

ACGGTGACGG CGCTACTCGC CGCGGCGGACG CGCGCAACGG CGCGATCCC CGCTTGCTCT	60
TGGGCACAGC CGGGGGGTGGC CGCGACGCTG CGCGCGGGCG CACCGGTACT CGCGGTGGCG	120
GCTTETGGGG CACCGAGGGC GACGGGGGCA CGCGCGGGCG TGCGGGGCTG TTAAATGGCG	180
CGGGCGCGCG CGCGACGCGT AGGACTGGCG CGCGCGGGCG TGCGCGTGTG CACGGGGCG	240
GCACGCGCG CGCGCGCGCG CGCGCGCGCA ACTGGCGGGCG CGGGGGGTAA CGCGCGCGTC	300
TGTTCTGGCG CGGGGGCGACG CGCGCGACCG CGCGATAKCG CGCGCGATCG CGTGGGGCG	360
CGACGGCGTT CGACGGCGACG ATGGCGGGCG TGCGGTGGTAC CGCTGCG	420

(2) INFORMATION FOR SEQ ID NO:174:

- (1) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 488 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GTGCGCTCG CGCGATCGCGC TGCGGGCGAA CGCGATCGCG CGTGTCAACG AAGAACATCG	60
TGGACCGCGCG CGCGCGGACG AACCGCGCTG CGTGGGGCGC TTGCGAAGCG CGCAGCGGGC	120

```

ACAGGCGTCCTGGGGCC TCGAGGCCA AGACCTGCT GTGGCGGGTG TAACTCGCG 186
GGATGCCCGC CTGGGGCGCA GATTCGGGC AGGGCGCGCG GTCTTGTGA TGTCTGAGA 246
TCACCGGAT STCTGGCGG ACCGGCGGGG GCGCGGAA GCTGGGGCGG CTTGGCGATA 306
GCGCGCGG GTGGGGCGG AGGTGGTGGG GGATGGTGGCG GCGGAGCGGT CGGGCGCGAC 366
GGCGGAAAG CGAACCGTCA CGCGGCTGG TCCGGCTGGG ATTCGGCTTG CGCTGGTGGG 426
CGATTTGCA CGCGGATGGC CGAGCGCGT ACAGGGCGG CGGGCGG 486

```

(b) INFORMATION FROM SIC ID NO. 3251

- (1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

exit SEQUENCE DESCRIPTION: SEQ ID NO:125;

```

GGCTGTTAACCG GGGGGCGAGGC TGGCATCGC GGCGGCGGGG AGCGAGGGGC CGAATGGGCGC 68
GAGGCCAATTC CTAAACGGCGC AAACCGCCGG AGCGGGGCGTA CGCGCTGCGA CGCTGCGGAC 124
GGCGGGCGCGC CGGGCGATGG CGGGCGCGCG AGCGAAGCGC AGCGGGCGCG CGTACACCGAC 180
GGCGGCAAGG GCGAGCGGGG CGAGCGGGGC AGCGGGCGC 216

```

INFORMATION FOR SGD ID: 100-126:

- (i) SEQUENCE CHARACTERISTICS:

 - (A) LENGTH: 494 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:176:

```

TACGTGGCGC GAGGCGCGCA AACGCCCGCA CCGTGCGACG GGCGGCTGAGC GCGTGCAGCG 60
CAACAGTTCG CTGACCCGAG CGCGGCGACGG CGCGTGCACCG CGCGCGCGCG CGCGCGCGCGCG 120
CAGCGGTTTT TTGCGGGCGA AGGGUGUCCG CGCGCGCGAC CGCGCGCGCG CGCGCGCGCG 180
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 240
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 300
CGATGGCGCG CGCTGCGACG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 360
CGACGCGCG CGCTGCGACG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 420

```

ATGCCGATCA GGCGCGTACT GCGCGCAGCA CCTGGGACAT CACCCGAGTT AGCGGGAGGC	480
ATTCCTGTAT CNCC	494

(2) INFORMATION FOR SEQ ID NO:177:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 220 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(3) SEQUENCE DESCRIPTION: SEQ ID NO:177:

GGGGGGGTGC TCGGGGGGGC CGCGCTTCTCA GCUCGGGGAGG CGCGGGGGAT GCGCTTGCGG	60
TTCGGCGCAC CGGGGGGGAG GGTGGGGCTG CGCGCTCGCG AGCGGGGGCG CGCGAUSCGC	120
CGGGCAAGCAC AGGTGTAACC GGTGGTACCG GGTGGCGCTG CGGGGGGGCG CGCGGTGGCG	180
CGGGGGGGCG CGGGGGGGAT GCGGGGGGCA TCAACGGCTC	220

(2) INFORMATION FOR SEQ ID NO:178:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 369 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(3) SEQUENCE DESCRIPTION: SEQ ID NO:178:

ATGGGGGCAA CGGGGGGGCG 66CGCGTGTG GGGGGGGCG CGACTACAT TTCCAAACGGC	60
GGGGCGGGTG GTGGGGGGCG CGAAGGGGGC CAAGGGGGCC TGGGGGGGCG AAGGGGGCG	120
TGATCGGGCT AGCCCGACCG CGGAGGGCGG ATCCGAGCG CGAGGATGG CGGTGGCTTG	180
CGGGGGTGGG CGGGGGGGCG ATCGCGTACG CTGACCGGGG CGGGGGCGATA CGGGGGGGCG	240
AGGGGGTGT CGGGGGGGCGT GCTAACCGGG CGGGGGGGCG ATACGGTGTG ACACGGTGTG CGGGGGGGCG	300
GGGGACTACG TCCGGGGCTG AGCATGGACA GGGGGGGCGA GTTGTGTGG ATGGGGATCGA	360
GGGGGGTGTG CGGGGGGGCG CGGGGGGGCG	396

(2) INFORMATION FOR SEQ ID NO:179:

- (1) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 460 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:179:

GCAAAGGCGG CACGGGGCGG GCGGGCATGA AGCGCTCGA CGCGCTGCTA GGCGGCCAAAG	60
ACGGCGCGCA AGCGGGCAGC GCGTGTCTGC GCGCGCAKGC CGGGGGGGCG CGCGCGCGCT	120
TGCGTCAGG CGCGACGCG: AAGCGCGCTA AGCGCGCTGA CGCGGGGTC CGCGCGAAACG	180
CGGGAAACCG CGCGAACCGC CGCGCAACA CGCGACCGCC CGCGCGGCG ACCCGAACCG	240
CGCACCGCGG CGCGCGCGG CGCGCGCGG CGCGCGCGG CGCGCGCGG CGCGCGCGG	300
CGCACCGCGG CGCACAAAGCG ACGGCGCGCR ACGGCGCGAC CGCGCGCAA CGCGCGCGCG	360
CGCGCGACCG CGCGACGCG CGCGACGCG CGCGCGCGG	420

(2) INFORMATION FOR SEQ ID NO:180:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 538 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:180:

GGCGACGGCG CGCGACGGCG CGACGGCGCG ATTCGGCGCG ATTCGGCGCG CGCGACGGCG	60
ACGGCGCGCG CGCGACGGCG CGCGCGCGG CGCGACGGCG CGCGACGGCG CGCGCGCGCG	120
CGACGGCGCG ACGGCGCGCG CGCGACGGCG CGCGACGGCG CGCGCGCGCG CGCGCGCGCG	180
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	240
CACCGCGCGT GACGGCGCGA CGCGCGCGA CGCGCGCGAT AACACCGCGA AGATGACTGC	300
CGACGGCGCG CGCGCGCGCG CGACGGCGCG CGACGGCGCG TTTCGGCGCG CGCGCGCGCG	360
CGCGCGCGCG CGCTTGACCGT CTGGGGCGCA CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	420
CGATGGCGCG AACGGCGCGA TGGCGCGCA CGCGCGCGTC ACTGGCGCG CGCGCGCGCG	480
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG	538

(2) INFORMATION FOR SEQ ID NO:181:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 239 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

EX-12 SECURITY INFORMATION: GRC III NOV 1984

GAGGCGCTTG	TACCCGGCGG	CAGCTTGTCA	GCCTGGGGGG	CGGGGGGGAT	GGCGTTGGGG	60
TTGGCGGCCAC	CGCGCGCCAG	GGTTGGGGGTG	CGGGTTCGGG	AGCGGGGGCG	GGCGAAGCGC	126
CGGCCAGCAC	AGGTGTATACG	CGTGCTACCG	ACTTCGCTGG	CGGGGGGGCG	GGCGTCGCGG	186
GGCAACCGCG	TAAGGGCAATT	GGGGGGGCGCA	TURACGGCTC	CGGTGTGCG	GGGGGGCGCG	239

(2) INFORMATION FOR SPC IN NO. 102.

(B) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 985 base pair
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 182;

AGGACAGCTTA CGGGTGGCGC CGGGTTGGCG GGGGGGGGGG GGGGGGGGGG
 60
 GGCAACAGCGG GTGTGCGGGG CGAATACGGG TCGGGGGGGG CGGGGGGGGC AGGGGGGAGG
 120
 GGGGGCCACCG GAGUTGCGGA CGGGTGGCGG CGGGGGGGGG CGGGGGGGGG
 180
 CGTGGGGGGGGG CGGGGGGGGG TGCGGGGGGG CGGGGGGGGG CGGGGGGGGG
 240
 ACGGGGGGCA CGGGGGGGGGT TCTCTGGGGGG AGCGGGTGGGG CGGGGGGGGG
 300
 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG AGCGGGTGGGG CGGGGGGGGG
 360
 TACGGGGGGGG AAGGGGGGGCA AGGGGGGGGG AGGGGGGGGG CGGGGGGGGG CGGGGGGGGG
 420
 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG
 480
 GGTCTCGGGGG AGAGGGGGGG CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG
 540
 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG AGGGGGGGGG AGGGGGGGGG
 600
 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG AGGGGGGGGG AGGGGGGGGG
 660
 CGGGGGGGGG CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG AGGGGGGGGG
 720
 TGGGGCTGGG CGGGGGGGGG AAAGGGGGGG TAAACAGGGGG CGGGGGGGGG
 780
 CGTGGGGGG CGGGGGGGGG AGGGGGGGGG AGGGGGGGGG CGGGGGGGGG
 840
 CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG
 900
 AGGGGGGGGG CGGGGGGGGG ATGGGGGGGG CGGGGGGGGG AGGGGGGGGG
 960
 CGGGGGGGGG CGGGGGGGGG CGGGGGGGGG AGGGGGGGGG CGGGGGGGGG
 998

(3) INFORMATION FOR SEC 12 NM 163:

- (i) SEQUENCE CHARACTERISTICS:
- LENGTH: 2138 base pairs
 - TYPE: nucleic acid
 - STRANDEDNESS: single
 - TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:1B3:

CCGCGAGG	AATGGTACCC	CCTGCATTCG	GCAGCTGODD	ATTCAACGGG	TTCGCCACCC	60
CGAGGAAAG	CGTAGGAGA	TGCGCGTCG	GAAGTAGGGC	GATCGCTTCG	CGATGCGCGC	120
ATGAAACGGC	GCATCAAT	TATGTGAGGA	ACGTTTGACT	TTAGGGACCA	TATGGGAT	180
AGCACTTAAG	AGGATHATTC	GATATGACG	ACTCGACAG	CTTCAGCTTC	GATGAGCAAG	240
AGATTTTAA	CADGCCAAC	GGGTGGAGG	CGCGATGCG	GGAGGCCCG	ACTGATGTCG	300
CCATCACCG	GTGCGACCT	ACGGGCGTA	AAACGCCCG	CTTACACGTC	TTATTGTCGG	360
CGGACACAT	GGCGGAAAC	CTGGCGGCCU	GTGCGAAAGA	GGGGAGCGT	CTGGCGACCT	420
CGCTGCGAA	CTGGGGCAAG	GGCTATGCG	AGTTTGATCA	GGAGGCTTCG	ACGGCGCTCG	480
ACGACGACG	CGGAGAACCT	CTGGCGCG	AACTGGCG	GGCGGTGCG	GGGGACAGTT	540
CGGGCGACAT	AAACGATACG	CGAGGCTTCG	CTACGGCG	TTAACCGAAC	TTCAAGKTC	600
TCGAAGAAC	GGCAAGGG	CTGAGAACG	GGGACCAAGG	CTCATGCTTC	GGGGCGTTG	660
CGGATGCGT	GGACGCTTC	AACTTGACG	TCGAGGCGA	CTTCAAGCGG	TTGGGGGGGT	720
TTGCGACCT	GGAGGCGT	GGCGCTACG	CTTGCGAGG	TTGCGCTGCT	CTACACCGCC	780
AAATGAACT	CTCATGCGC	AATTGACCG	CTGGATGTCG	CTAGATGTCG	AAATATGTCG	840
CGGACATGCA	CTTGGGGCT	AGGGGGGAACT	ATCGGACTTA	TCGAGAACAT	CTGGGGCTCG	900
AAACGCTTAA	CGGGGAAAC	CTTGGGGCG	GGGACCAAG	TCGCGGCTG	TCGGGGCGAT	960
ATCGGCGAA	CTGGGGCG	GGTGGTAACT	ATTCACACCA	CTACACCGCC	CTGGGACCGG	1020
TCACCGCGAC	GAAGGCTCG	CGGGCGCTCA	AGATGACCG	GGGGCGCGT	GGGACAGACG	1080
AGGGATTTAA	CGCTGGCTTC	CTGATGCGCG	CGCTGGACGG	CTCCGGTGTG	ACTGGCGGTA	1140
CGGGGATCG	AGGGCGACG	ATGTTTCGCG	CTACGGGATC	GGGGCGCTGT	GGGGCGCGCG	1200
CTGACACCG	GGGGCGCTG	ACCTGGCGTC	GGGGGAGACG	CTGAGGGCTG	TCGGGGACCG	1260
TCGGGGTCA	AGGGGGATCG	CTGGGGCTTC	CTGGGGAGCG	GGGGGGCGT	TCGGGGCGGT	1320
TCGGGGATCG	GGGGGGGGCG	GGGGGGATCG	TCGGGGCGCG	TCGGGGCTGT	GGGGGGGGGT	1380
GGTGGGGCCA	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGGT	GGGGGGGGGT	1440
GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGCG	GGGGGGGGGT	GGGGGGGGGT	1500

ACCGGCGCTT CTAGAACGGG GATCGGGCAT GGACCCAGGC CGTCATTGGT AACCGCTGGC	1560
GGCAGGAGAG TAAGGAGTGT AAGTGAGCAT GGACGGATTG GACCCGCGTA TCGGCCGGGC	1525
GTTOACCGTG GGGCGGGGT TTGAGTCGGC CCTAGAACGG AGCTTCATACT AGTGTGACAA	1590
CGGATCCTTC CGGGCCACCG AGGAGGCGA GACCGCGTA CGTACGGATCA ATGGCGACCA	1745
GTGGCTTACCG GGCTCGCGA TCGAAGATGG TTGGCTGAAG AGGCTGGGT CGCAGGGGT	1800
GGCTGCGGGG CTGAAUGGG CGCTGCGAC TCGCGGGGC CGGGCTGGC CGTATACGCA	1860
CUCGGGGGGC GAGCGCTGA CGCGTGCCTT ATCGGCGATG TCCGGCGCA TGAAGTAAAG	1920
ATGGCGCTAA CGCCATGGT GCCTGGTTAG CGACTACCGA CGUATGAGC CGCGCATGCG	1980
GGTCATTCGG CGCGCGCCCG ACGGCGTAG AGCGATTTT CTATGTTTAC ACATGGATCG	2040
CGCGGTTGG GAGGGCGCCA TAGTGCTGT CGCGATATT CGCGCGCTTA CGTGCGCTTA	2100
CGTGGGTTA CGCTGGTAA TTATGAGTC CGTACCA	2135

(7) INFORMATION FOR SEQ ID NO:194:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 460 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Met Thr Glu Ser Gin Thr Val Thr Val Asp Gin Glu Ile Leu Asn			
1	5	10	15
Arg Ala Asn Glu Val Gin Ala Pro Met Ala Asp Pro Pro Thr Asp Val			
20	25	30	
Fro Ile Thr Pro Cys Glu Ieu Thr Ala Ala Lys Asn Ala Glu Gin Glu			
35	40	45	
Leu Val Leu Ser Ala Asp Asp Met Arg Glu Tyr Leu Ala Ala Gly Ala			
50	55	60	
Lys Glu Arg Glu Arg Ieu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala			
65	70	75	80
Tyr Gly Glu Val Asp Glu Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly			
85	90	95	
Glu Glu Thr Val Gin Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser			
100	105	110	
Ser Ala Glu Ieu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro			
115	120	125	
Asn Phe Met Asp Ieu Lys Glu Ala Ala Arg Iys Ieu Glu Thr Gly Asp			
130	135	140	

Gln Gly Ala Ser Leu Ala Asp Phe Ala Asp Gly Trp Asn Thr Phe Asn
 145 150 155 160
 Leu Thr Leu Gin Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp
 165 170 175 180
 Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ser Leu Asp Gin Gin Arg
 185 185 190
 Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Asp Met Ala Lys Gln
 195 200 205
 Ala Gln Tyr Val Ala Gln Leu His Val Trp Ala Arg Arg Gln His Pro
 210 215 220
 Thr Tyr Glu Asp Ile Val Gly Leu Gln Arg Leu Tyr Ala Glu Asn Pro
 225 230 235 240
 Ser Ala Arg Asp Gln Ile Leu Pro Val Tyr Ala Gln Tyr Gln Gln Arg
 245 250 255
 Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys Ala Ala Leu Glu Pro
 260 265 270
 Val Asn Pro Pro Lys Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro
 275 280 285
 Pro Pro Gln Gln Gln Gly Leu Ile Pro Gly Phe Ieu Met Pro Pro Ser
 290 295 300
 Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Pro Met
 305 310 315 320
 Val Pro Pro Thr Gly Ser Pro Gly Gly Leu Pro Ala Asp Thr Ala
 325 330 335
 Ala Gln Leu Thr Ser Ala Gly Arg Gln Ala Ala Leu Ser Gly Asp
 340 345 350
 Val Ala Val Lys Ala Ala Ser Leu Gly Gly Gly Gly Gly Val
 355 360 365
 Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Gln Ser Val Arg
 370 375 380
 Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly
 385 390 395 400
 Gly Gly Ala Ala Leu Gly Gly Gly Met Gly Met Pro Met Gly Ala
 405 410 415
 Ala His Gln Gly Gln Gly Gly Ala Lys Ser Ieu Gly Ser Gln Gln Gln
 420 425 430
 Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Gln Ala Val Ile
 435 440 445
 Gly Asn Arg Arg Arg Gln Asp Ser Lys Glu Ser Lys
 450 455 460

(2) INFORMATION FOR SEQ ID NO:185:

16.2 SEQUENCER CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

Table 3: Sequence description: NCO vs. NC

Ala Gly Asn Val Thr Ser Ala Ser Gly Pro His Arg Phe Gly Ala Pro
 5 10 15
 Ile Arg Gly Ser Gln Arg Arg Arg His Pro Ala Ala Ser Thr Ala
 20 25 30
 Thr Glu Arg Cys Arg The Asp Arg His Val Ala Arg Gln Arg Cys Gly
 35 40 45
 Phe Pro Phe Ser Arg Arg Gln Leu Arg Arg Arg Val Ser Arg Glu Ala
 50 55 60
 Thr Thr Arg Arg Ser Gly Arg Arg Asn His Arg Cys Gly Trp His Pro
 65 70 75 80
 Gly Thr Gly Ser His Thr Gly Ala Val Arg Arg Arg His Gln Glu Ala
 85 90 95
 Arg Asp Gln Ser Leu Leu Leu Arg Arg Arg Gly Arg Val Asp Leu Asp
 100 105 110
 Gly Gly Gly Arg Ileu Arg Arg Val Tyr Arg Phe Gln Gly Cys Leu Val
 115 120 125
 Val Val Phe Gly Gln His Leu Leu Arg Pro Leu Leu Ile Leu Arg Val
 130 135 140
 His Arg Gln Asn Leu Val Ala Gly Arg Arg Val Phe Arg Val Lys Pro
 145 150 155 160
 Phe Gln Pro Asp Tyr Val Phe Ile Ser Arg Met Phe Pro Pro Ser Pro
 165 170 175
 His Val Gln Leu Arg Asp Ile Leu Ser Leu Leu Gly Gln Arg Ser Ala
 180 185 190
 Gln Phe Gly His Val Gln Tyr Pro Leu Pro Leu Leu Ile Gln Arg Ser
 195 200 205
 Leu Ala Ser Gly Ser Arg Ile Ala Phe Pro Val Val Lys Pro Pro Glu
 210 215 220
 Pro Leu Asp Val Ala Leu Gln Arg Gln Val Gln Ser Val Pro Pro Ile
 225 230 235 240
 Arg Lys Val Arg Gln Arg Cys Ala Leu Val Ala Arg Phe Gln Leu Pro
 245 250 255
 Cys Arg Phe Phe Glu Val Ile His Gln Val Gln Phe Thr Gly Arg Gly His

260	265	270
Pro Arg Arg Ile Gly		
275		

(2) INFORMATION FOR SEQ ID NO:186:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 192 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:

Arg Val Ala Ala Ser Phe Ile Asp Trp Leu Asp Ser Pro Asp Ser Pro	15																																																											
1	5	10	15	Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly		20	25	30	Ala Glu Ser Ser Ala Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg		35	40	45	Tyr Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr		50	55	60	Ser Arg Asp Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg		65	70	75	80	Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg		85	90	95	Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser		100	105	110	Gly Arg Pro Ile Ser Ser Phe Pro Ile Val Arg Ser Arg Ser Cys Val		115	120	125	Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Ile Asn Arg		130	135	140	Ala Arg Ser Ser Ser Ile Ala Ala Arg Cys Arg Thr Ile Gly Thr Phe		145	150	155	160	Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro		165	170	175	His Val Thr Lys Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly		180	185	190
10	15																																																											
Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly																																																												
20	25	30																																																										
Ala Glu Ser Ser Ala Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg																																																												
35	40	45																																																										
Tyr Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr																																																												
50	55	60																																																										
Ser Arg Asp Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg																																																												
65	70	75	80																																																									
Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg																																																												
85	90	95																																																										
Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser																																																												
100	105	110																																																										
Gly Arg Pro Ile Ser Ser Phe Pro Ile Val Arg Ser Arg Ser Cys Val																																																												
115	120	125																																																										
Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Ile Asn Arg																																																												
130	135	140																																																										
Ala Arg Ser Ser Ser Ile Ala Ala Arg Cys Arg Thr Ile Gly Thr Phe																																																												
145	150	155	160																																																									
Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro																																																												
165	170	175																																																										
His Val Thr Lys Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly																																																												
180	185	190																																																										

(2) INFORMATION FOR SEQ ID NO:187:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 196 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:

(D) TOPOLOGY: linear

(xx) SEQUENCE DESCRIPTION: SEQ ID NO:187:

Gln Gln Arg Pro Glu Met Cys Glu Arg Val Ser Glu Ile Gln Pro Arg
 1 5 10 15

Thr Gln Phe Phe Asn Arg Cys Ala Leu Pro His Tyr Trp His Phe Pro
 20 25 30

Ala Val Ala Val Phe Ser Lys His Ala Ser Leu Asp Glu Leu Ala Pro
 35 40 45

Arg Asn Pro Arg Arg Ser Ser Arg Arg Asp Ala Glu Asp Arg Arg Val
 50 55 60

Ile Phe Ala Ala Thr Ile Val Ala Val Asp Pro Pro Leu Arg Gly Ala
 65 70 75 80

Gly Gly Glu Ala Asp Glu Leu Ile Asp Leu Gly Val Cys Arg Arg Gln
 85 90 95

Ala Gly Arg Val Arg Arg Gly Gln Glu Leu His His Arg His Arg Gln
 100 105 110

Gln Gly Ala Ala Pro Asp Leu Arg Arg Arg Arg Arg His Arg Arg Val
 115 120 125

Gln Gln His Arg Arg Leu Gln Arg Val Arg Gln Leu Arg Arg Tyr Val
 130 135 140

Gln Thr Ala His His Arg Arg Phe Ala Arg Thr Asp Arg Val Arg His
 145 150 155 160

His Val Arg Gly Pro Ser Asn His Arg Arg Arg Arg Val Tyr Arg Gly
 165 170 175

Arg His Ser Gly Ala Gly Gly Cys Pro Ala Gly Gly Ala Gly Ser Val
 180 185 190

Gly Gly Ser Ala
 195

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xx) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Val Arg Cys Gly Thr Leu Val Pro Val Pro Met Val Glu Phe Leu Thr

1	5	10	15
Ser Thr Asn Ala Pro Ser Leu Pro Ser	Ile Tyr Ala Glu Val Asp Lys		
20	25		30
Leu Ile Gly Leu Pro Ala Gly Thr Ala Lys Arg Trp Ile Asn Gly Tyr			
35	40	45	
Gly Arg Gly Gly Lys Asp His Pro Pro Ile Leu Arg Val Thr Pro Gly			
50	55	60	
Ala Thr Pro Trp Val Thr Trp Gly Glu Phe Val Glu Thr Arg Met Leu			
65	70	75	80
Ala Glu Tyr Arg Asp Arg Arg Lys Val Pro Ile Val Arg Glu Arg Ala			
95	90	95	
Ala Ile Glu Glu Leu Arg Ala Arg Phe Asn Leu Arg Tyr Pro Leu Ala			
100	105	110	
His Leu Arg Pro Phe Leu Ser Thr His Glu Arg Asp Leu Thr Met Gly			
115	120	125	
Gly Glu Glu Ile Gly Leu Pro Asp Ala Glu Val Thr Ile Arg Thr Gly			
130	135	140	
Gln Ala Leu Leu Gly Asp Ala Arg Trp Leu Ala Ser Leu Val Pro Asn			
145	150	155	160
Ser Ala Arg Gly Ala Thr Leu Arg Arg Leu Gly Ile Thr Asp Val Ala			
165	170	175	
Asp Leu Arg Ser Ser Arg Glu Val Ala Arg Arg Gly Pro Gly Arg Val			
180	185	190	
Pro Asp Gly Ile Asp Val His Leu Leu Pro Phe PRO Asp Leu Ala Asp			
195	200	205	
Asp Asp Ala Asp Asp Ser Ala Pro His Glu Thr Ala Phe Lys Arg Leu			
210	215	220	
Leu Thr Asn Asp Gly Ser Asn Gly Glu Ser Gly Glu Ser Ser Glu Ser			
225	230	235	240
Ile Asn Asp Ala Ala Thr Arg Tyr Met Thr Asp Glu Tyr Arg Glu Phe			
245	250	255	
Pro Thr Arg Asn Gly Ala Glu Arg Ala Leu His Arg Val Val Thr Leu			
260	265	270	
Leu Ala Ala Gly Arg Phe Val Leu Thr His Cys Phe Ala Gly Lys Asp			
275	280	285	
Arg Thr Gly Phe Val Val Ala Leu Val Leu Glu Ala Val Gly Leu Asp			
290	295	300	
Arg Asp Val Ile Val Ala Asp			
305	310		

(2) INFORMATION FOR SEQ ID NO:189:

(iii) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2072 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(E) SEQUENCE DESCRIPTION: SEQ ID NO:185:

CTGTGTCCGA	TTCGGGCGCG	GTCAGGACGC	CCAGGGGCC	GTTGGGCTTA	GTCATCGAGG	60
CATTGGCGA	CGGGCTGCC	GGCAAGGCTA	NGCAAACTAA	CCACACGTC	AACAGGCTGT	120
CGGGGGCTT	GAAGGCTTG	ATGAGGGGG	GGGGGACTT	CTTGGGCTTG	GTACGACCC	180
TGGCCCTATT	CGTCAACGGG	CTACATCAGG	AGGACGACA	GTTGGTGGG	TTGACACAGA	240
ACCTTGGGAA	GTTCAACGAC	AGGTTAACCC	ACTCGGTTGC	GGACGTGTCG	AACGCCATCC	300
ACGATTCGAA	CAGCTTCTC	GGGTCGCGC	GGCGTTTCTT	GGCGAGGAG	GGCGAGGTGC	360
TGACCCATGA	CTGCAATAT	CTDGCGACCG	TWACCAACAC	GTTCTGCG	CCCGATCCGT	420
TGGATGCTT	GGGAGGCGTC	CTGCAATAT	TGGAGACCT	GGGAGGCGAC	ATTAACAGAC	480
TTTACCATCC	GACACACCT	GGGGGGGTGT	GGCTTCCCGC	GTTCACGAA	TTGGCCAAAC	540
CCATGCGTT	CATCTCGACG	TGATTTCAGG	GGGTACGCC	GGCGTTTAT	CAAGAGTGG	600
CCGAACTCTG	TGGCGCATAT	CTGGCGCGAG	TCTTGATGC	GATCAAGGTC	AACACTTTC	660
GGTTCGGCTT	GAACGTTGGC	AGGCGGCGCT	GGACACTCC	TAAGAGATC	GGGTACTCGG	720
ACCCCGGTT	GGCAACGCGG	AAACGGTACA	ARGAACACAC	GGTGGCGCG	ATCTGGGTCG	780
CGTGTACGCG	GTTCGGTACG	GGCAACGCG	AAACGGTACA	GGTGGCGCG	GGCGGGGTC	840
AAAGGGTTCA	GGTGGGACCG	KTACGGCAGG	GGTTCGGTAC	GGCGGAGTCG	CTGGCCGAAAC	900
TCATCGCTGG	TGGCGATAT	GGCGCTCGT	GGTCAGGCGT	GGAAACCGG	GGGGACCC	960
CGAAATCGTA	CGGGGAGGAC	CCCCTGCTCG	GGGGGATGG	TTTACAGGG	CGGGGGTTCG	1020
GGATAACCGC	GGCGCTCTCT	GGGGGGGG	TATGCGCG	GGCGGTGGCA	GGGGGGTTGG	1080
GGCGGATGTT	GGTGGCGAGA	TGTCGCGG	GGCGGTGGCA	GGGGGGGG	GGGGGGGG	1140
TGGTGTGTC	GGCGGCGGG	CTGGGGAGCT	TCTGGTTCGA	GGTGTGTCAT	AGGGGGGGGG	1200
GGGGGGGGAT	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1260
GGCGGATGTT	GGCGGCGGG	TGTCGCGG	GGCGGTGGCA	GGGGGGGG	GGGGGGGG	1320
TGGGAAACG	AGGGGGGG	GGGGGGGG	TGAGGATGAG	GGGGGGGG	GGGGGGGG	1380
TTGGGTTCTT	GGTGGCGGTC	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1440
AAATGGGGGT	GGATATGTC	GGGAGGGGGC	TGGTGGCGAT	GGGGGGGG	GGGGGGGG	1500
GGGGGGGGAT	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	GGGGGGGG	1560

CTGGCGGGCTT GGCACACTTC SCCTTTGCTG TGCGCGGCCA AGCGGAACTAC TTACCCAGTC	1620
TGCCGCGCGG CCTGGCAATC ATGGGATGGG GGTGGGGGCTG CTXAAKRTG DGAATGTCCTG	1660
GGGGCGGAGT GCAGACCTTG GCGCACACAU AGATGGCTCG CCGTGGAGG CTGATCAGCG	1740
TCAACAGCGA CGTGGGGCGGT TCGATAGGGG CGCGCACTGAT GTGGGTGCTG CTGCGCTNCG	1800
AATTCRATCA CGCGAAATC ATGGCTCTG CAAGAANCG CGCACTGACG CGAGAGAGTC	1860
GGCGGGGGG GGGGGGGCGC GTGAGGCTT CCTGGCTTACG CGCGCAAAACG AACGGGGGGG	1910
CCCAACTGCT GCATGACCTT TCGCGCGCGT AGCGGTGTT ATTGGTGATA CGGAGCGCG	1980
TAGTGCTG GACGTGATG CGCGGGCAT TCGCGCGAA AGCGGAGGT AGTCATCGAA	2040
GAGCACCGTT GGTATCGCGA TSACGCTCG TT	2072

(2) INFORMATION FOR SEQ ID NO:180:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1923 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:180:

TGACCCCCGA GAATCTTTG CTGACAGATT TGGACATGCA CTGGCTGCG ATGGTCGAGA	60
TGCGCTTCA GACCGAGGAC AATGACCGCG TCAAGATCC CGACAGACG CTGCGCGCTG	120
TGGTTCGCGT CGTGTGAGTT CTGGCTTAAG TCGGAGCTG CGAGGAGAGA AACCGGGAGG	180
CGGCTGAGG CGTGGGGGGG AGATTTGAGT CGGGAAGCC CGATGCGCA CGAGGAGAG	240
CGTGGCTTC ACCACATCG CGAGCTGAG ACACCGCTCG TCGTCTTCG CGCTCACCA	300
GATTTGCGT TCGGGGGCTT CGCGCACTG TTCCGACAC AGCGGAGAC CGTGGCGAA	360
CGATGCTGAT CGGGGGACCA CGATGCTGAT CGCGCGCTGC CTGACAAATG CGCGCGCTG	420
CGATTTGGCG CGGCGGAAATT CGGGCGATT CGCGAACGCG AGCGATCG CGCGCGCGCG	480
CGCGCGCTG CGATTTGGCG CGCGCGGAT CGCGCGGAC AGCGGAGAGA TGGCGGGAG	540
GAACAGCGT AGCGGATCA CGACCGAGG ATTCGACAC CGATCGCTG CGTACCGCGTC	600
CGCCCGGCTT CGTCGATCGCG CGCGATGATCG ATGGGGGGGT TTAAGCTAAC AGCTTTGCGG	660
GGACCGGGGG TGCGACCGCG CGATGCTGC CGCGCGGAGAC CGGGAGGTC CGCGCGCG	720
TGCGCGCG CGATGCTGC ACCATCGGGG TCGCGCGCG CGCGCGCTG CGCGCGCG	780
AACCGACAGG GTTCTCTAG CGACCGAGTC AACCTCGACG CGCGCGCTT CGTACGTTTC	840
CGCGCGCGC AGGTCTCGCG CGACGCGCTT CGCGCGCGCG CGTGGGGCG CGTGGCGAG	900

CAGGGGATCG TAACTTGCSC CACCGCTGAC ATCGTCTTCG CGGAGGTGTGT CGGTCAAGCC	960
CGGATATGAG CAGGGATCGA ATGCCAACATA ATCTGCTGGG CTGAGTTGCG CCTAGTACG	1020
GTGGACGCCA ACGGGGCGAA TACAUATGCGC CCGTGGTACG CGGGTCAAAGA CGGAATAGCT	1080
TTCCACAGGC CGCTGUGCGA TCAAGATGACG GCAACGGTTC AGCGCGCGCA CGGGCGCCCTC	1140
GTGCTCTTG CGCCACGTTG CGAAATCGGC AAGCGGAGCG CTGGGTGTCG GTGGCGATCG	1200
CGCGGTGTCG GTGGCGAGGT TTCCGGAACT ATTCGCTCGG TCAACGGGGG CGGGCGAGCT	1260
TCTGGCGCTG CGAGGAGAAC CGAGCGCTTC CGACGAGTT CGACACGAGT CGGGGCGCGC	1320
TCAATCTCGA TCGCCCCCTG GCGCTGGTG KCTCCACCTG GGTGCTTCCC GGCGCAGGLA	1380
AGCGCTCGG CGATCGGTT CGGATACCC AGCGCTCTG CGACACGATG CGCTCTTCGG	1440
ATGGTAAAGA ATTAAACATT CGACGCTCA TAGCGCTGTC CTGGCGCGA TGTCGGGAGG	1500
ATCGCGTAGU CTATCGAACG ATTCGTTGGG AAATGGCTGA CGGAGCGTTCG GTTGGGGGTTG	1560
ATGGGTGTCG ATCCCGGTTG GAGCGGATCG GGGCTGTCG TCATCGGAG TGCGCGCTGGT	1620
CGGGCGCTCA CGCGCTCGA TCGACGCTG CGGGCACAC CGCTCGATGC CGCTCTGGCG	1680
CGGGCGCTGT CGATGCCGTC GAGCACTCG CGGACACCG CGTACGGGAG CGCTCTGGCG	1740
GTGGTGGCTA TCGAACGGGT GTTCTCTCG CTCAAAGTGA CGCGCTGTTG GCGCACCGCG	1800
CGGGCGGGG GCGTCGATCG CGCTGGGGCG CGCAACGCTG CGTGGACGCT CGATTTGCGT	1860
ACCCCGAGCG AGGTCAGAACG CGGGCGTGT CGCAACGTT CGCGACGACG CGCTCGAGTC	1920
ACC	1923

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1055 base pairs
- (B) TYPE: nucleic acid
- (C) STRANGENESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

CTGGCTTCC AGTCAGACAGG CGGATATGAG ATCCGGATTC ATTTGGCGG CGCGCGCGAA	60
CGCGCGCGCA CGGACGCTGG ACCACCGGTT CGGATACCCG CGGAGCTG CGAAGCTGCT	120
CGGGATGTCG GTGGCGAGGT TTCTGAAAGA AAGCGGAGCG TATCGGCTCG CGTACGGGAG	180
AAAGACGCGA CGCGCTCGCA CGGGCGGATT CGGAGCGCA CGCGCTGAGCG CGGGCGCTGG	240
GGGGATGGTT CGAGCGTAAC CGTGGCGCTG CGCGAACCG CGGGCGCGC CGCTGGACCG	300
GCGTACGGCG CGGAATTTAG AAGACCGCGA ATTTGTCGCG CGGATGGTCA TGGATGTCG	360

ACTGCTTGT GGTTCGGTGT TAAACATTGA CTGTCGATGT GTTGAGAACC TGGCGTCAA	420
GGGGCAACG TGCCTTCTGCC TTCTTGCCG CTTTGGCCG CGCTTCGCT GTTTCGAA	460
CGGGTGTGCT CGGGATGCT CGGGCGGGTT TCGGTTCCGC TCAAGCTGGG ATCGCTGCTC	540
GGACCGATG ATTATGACCA CTGGGGCTGT CGCGCGTGGG AGGGGGCGGA TUTCGATGTC	600
CAGGGGGGGG AAAGGGGGGA CGCAGGGCC CGCTGGCATGG AGGAGTGGAA TGAGTGCGAG	660
GCGTGGACG ACTGGCTGGC GGAGAACGCT GAGCCCGCT TTGAGGCTCC ACGGGAGTAGC	720
ACGACGCTGA TTCCCGATTC TCGGGGGGGG CGCTAGGAGA GGAGGGCGAG ACTCGCTGTA	780
TITGACCACT GATCGGGCTG CTGGGGTGTG CGGGGGGGG CTATGACNN AGTCAATGTC	840
CATGACAACT TACAGGTTAT TGGTCAAGT TGAACAGGA GCGACGCAAC ATGGCAACAC	900
CGTGGATAC GGATCCGAC GGGATGGGGG ACATGGGGGG CGCTTITGAG GTGACCGCCC	960
AGACGGTGGG GGAGGAGCT CGGGGGATGT GGCGCTGGC GCGAACATC TGGGGGGGG	1020
GTGAGTGG CGGGGGGGG CGGACCTGGC TAGAC	1055

(2) INFORMATION FOR SEQ ID NO:192:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 359 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:192:

CCGGCTTGT GTGGCATAAC TCGGGGGGGG CGCTCTGAGC CGCACTGGCC GTGGGGCTTC	60
TCCGGCTGTA CCACGGGGAT CGGGAAACCA TGGGAGACCA CCTCGCAATG ATCCACCTCG	120
CGCACGGCTG CTACGGCGCA CGGGGGGGT TGGAGAGGG CGCTCTGAGC CTGGGTATAG	180
CGTGGGGGGC CGAGGGGGG GAGGTTAG TACTGGCCCA CGAGCTGGT ACGGGGGGCG	240
GGAGTTCA CGGTGGAGGT CGGGATGCTG CGGGGGGGT ATGGACCCG GAAACCGAG	300
TCTTCGGCA GGTCTCGGG CGGGGGGGC AGGACAAACG CGAGGGGGGG ATAGGTTAG	359

(2) INFORMATION FOR SEQ ID NO:193:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

ATCGGCTTC TGGCGCGCG CGCTCTTACG GCTCTCTTG CTCGCTGAA GTCGTGAAG	60
GATGCATTT CGCAGATTEC CGCCAGAGCA ARACACCGCG TACTCTTAGT CGGAGTCGGC	120
CGTAAAGTC CTCGAATAAC TCGGTACCGS GACCCCAA CGCGGCTCC TTCTCTTACG	180
TGCGCGAACG ACTTGAGGTT CGGGGACTTC TTGACCTCCA CGCGGATTCG TTGGAGTGGC	240
TGATGGGTC CGCGCGCTCG CGCGGATTCG CGCGCGAGCG CGGTGATTCG AACCGACTGG	300
CTGGCCTGGA AGGGGTGTCG TACGAGCTGT CTCCGATCGA GAGCTCTCC	360

(2) INFORMATION FOR SEQ ID NO:194:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 679 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Gly Ser Pro Lys Gly Pro Phe Gly Glu Val Ile Glu Ala Phe Ala Asp	
1 S 10 15	
Gly Ile Ala Gly Lys Gly Lys Glu Ile Asp Thr Thr Leu Asn Ser Leu	
20 29 30	
Ser Glu Ala Ieu Asn Ala Leu Asn Glu Gly Arg Gly Asp Phe Phe Ala	
35 40 45	
Val Val Arg Ser Ieu Ala Ieu Phe Val Asn Ala Ieu Ser Glu Asp Asp	
50 55 60	
Gln Gln Phe Val Ala Leu Asn Lys Asn Leu Ala Glu Phe Thr Asp Arg	
65 70 75 80	
Leu Thr His Ser Asp Ala Asp Leu Ser Asn Ala Ile Gln Gln Phe Asp	
85 90 95	
Ser Leu Leu Ala Val Ala Arg Pro Phe Phe Ala Lys Asn Arg Glu Val	
100 105 110	
Ieu Thr His Asp Val Asn Asn Leu Ala Thr Val Thr Thr Leu Leu	
115 120 125	
Gln Pro Asp Pro Leu Asp Gly Leu Glu Thr Val Leu His Ile Phe Pro	
130 135 140	
Thr Leu Ala Ala Asn Ile Asn Gln Leu Tyr His Pro Thr His Gly Gly	
145 150 155 160	
Val Val Ser Leu Ser Ala Phe Thr Asn Phe Ala Asn Pro Met Glu Phe	
165 170 175 175	

Ile Cys Ser Ser Ile Gin Ala Gly Ser Arg Leu Gly Tyr Glu Glu Ser
 180 185 190

Ala Glu Leu Cys Ala Gin Tyr Leu Ala Pro Val Leu Asp Ala Ile Lys
 195 200 205 210

Phe Asn Tyr Phe Pro Phe Gly Leu Asn Val Ala Ser Thr Ala Ser Thr
 210 215 220

Leu Pro Lys Glu Ile Ala Tyr Ser Glu Pro Arg Leu Gin Pro Pro Asn
 225 230 235 240

Gly Tyr Lys Asp Thr Thr Val Pro Gly Ile Trp Val Pro Asp Thr Pro
 245 250 255 260

Leu Ser His Arg Asn Thr Glu Pro Gly Trp Val Val Ala Pro Pro Gly Met
 265 265 270

Gln Gly Val Glu Val Gly Pro Ile Thr Glu Gly Leu Leu Thr Pro Glu
 275 280 285

Ser Leu Ala Glu Leu Met Gly Gly Pro Asp Ile Ala Pro Pro Pro Ser Ser
 290 295 300

Gly Leu Glu Thr Pro Pro Gly Pro Pro Asn Ala Tyr Asp Glu Tyr Pro
 305 310 315 320

Val Leu Pro Pro Ile Gly Leu Glu Ala Pro Glu Val Pro Ile Pro Pro
 325 330 335

Pro Pro Pro Gly Pro Asp Val Ile Pro Gly Pro Val Pro Pro Val Leu
 340 345 350

Ala Ala Ile Val Phe Pro Arg Asp Arg Pro Ala Ala Ser Glu Asn Thr
 355 360 365

Asp Tyr Met Gly Leu Leu Leu Leu Ser Pro Gly Leu Ala Thr Phe Leu
 370 375 380

The Gly Val Ser Ser Ser Pro Ala Arg Gly Thr Met Ala Asp Arg Asn His
 385 390 395 400

Val Leu Ile Pro Ala Ile Thr Thr Gly Leu Ala Leu Ile Ala Ala Phe Val
 405 410 415

Ala His Ser Trp Tyr Arg Thr Glu His Pro Leu Ile Asp Met Arg Leu
 420 425 430

The Gin Asn Arg Ala Val Ala Gin Ala Asn Met Thr Met Thr Val Leu
 435 440 445

Ser Leu Gly Ile Phe Gly Ser Phe Leu Leu Leu Pro Ser Tyr Leu Glu
 450 455 460

Gln Val Leu His Gin Ser Pro Met Gin Ser Gly Val His Ile Ile Pro
 465 470 475 480

Gin Gly Leu Gly Ala Met Leu Ala Met Pro Ile Ala Gly Ala Met Met
 485 490 495

Asp Arg Arg Gly Pro Ala Lys Ile Val Leu Val Gly Ile Met Leu Ile
 500 505 510

Ala Ala Gly Leu Gly Thr Phe Ala Phe Gly Val Ala Arg Gln Ala Asp
 515 520 525
 Tyr Leu Pro Ile Leu Pro Thr Gly Leu Ala Ile Met Gly Met Gly Met
 530 535 540
 Gly Cys Ser Met Met Pro Leu Ser Gly Ala Ala Val Gln Thr Leu Ala
 545 550 555 560
 Pro His Gln Ile Ala Arg Gly Ser Thr Leu Ile Ser Val Asn Gln Gln
 565 570 575
 Val Gly Gly Ser Ile Gly Thr Ala Leu Met Ser Val Val Leu Thr Tyr
 580 585 590
 Gln Phe Asn His Ser Glu Ile Ile Ala Thr Ala Lys Lys Val Ala Leu
 595 600 605
 Thr Pro Glu Ser Gly Ala Gly Arg Gly Ala Ala Val Asp Pro Ser Ser
 610 615 620
 Leu Pro Arg Glu Thr Asn Phe Ala Ala Gln Ile Leu His Asp Leu Ser
 625 630 635 640
 His Ala Tyr Ala Val Val Phe Val Ile Ala Thr Ala Leu Val Val Ser
 645 650 655
 Thr Leu Ile Pro Ala Ala Phe Leu Pro Lys Gln Gln Ala Ser His Arg
 660 665 670
 Arg Ala Pro Leu Leu Ser Ala
 675

(2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 125 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

Thr Pro Glu Lys Ser Phe Val Asp Asp Leu Asp Ile Asp Ser Leu Ser
 1 5 10 15
 Met Val Gln Ile Ala Val Gln Thr Glu Asp Lys Tyr Gly Val Lys Ile
 20 25 30
 Pro Asp Glu Asp Leu Ala Gly Leu Arg Thr Val Gly Asp Val Val Ala
 35 40 45
 Tyr Ile Gln Lys Leu Gln Gln Glu Asn Pro Glu Ala Glu Ala Leu
 50 55 60
 Arg Ala Lys Ile Gln Ser Glu Asn Pro Asp Ala Ala Arg Ala Asp Arg
 65 70 75 80

Cys	Val	Ser	Pro	Thr	Ser	Gln	Ala	Arg	Asp	Ala	Arg	Arg	Pro	Ile	Ala
						85			90					95	
Arg	Ser	Ala	Arg	Leu	Ala	Cys	Arg	Arg	Leu	Pro	Ala	Ser	Val	Pro	Thr
				100					105				110		
Thr	Arg	Arg	Asp	Phe	Arg	Gly	Arg								
				115					120						

(2) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 85 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:196:

Lys	Ala	Cys	Gln	Cys	His	Arg	Arg	Tyr	Asp	Val	Gly	Ile	Gln	Met	Arg
						5				10				15	
Gly	Pro	Ala	Gly	Pro	Val	Ala	Thr	Gln	Ser	Gly	Pro	Pro	Gly	Pro	Ser
					20			25					30		
Ile	Ala	Glu	Gly	Arg	Gin	Val	Arg	Ala	Gln	Gly	Ala	Gly	Phe	Leu	
		35						40				45			
Glu	Arg	Arg	Pro	Ala	Val	Ser	Gly	Ala	Leu	Pro	Pro	Asn	Asn	Ala	Ser
		50			55							60			
Pro	Gly	Ile	Arg	Ser	Arg	Ala	Ala	Asp	Ser	Gln	Arg	Ile	Ile	Ala	
		65				70				75			80		
Gly	Asp	Gly	Ser	Asp	Val	Thr	Val	Gly							
					85										

(2) INFORMATION FOR SEQ ID NO:197:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 119 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:197:

Ala	Ser	Leu	Leu	Ala	Tyr	Ser	Ala	Ala	Ala	Ser	Thr	Ile	Leu	Ala	
							5				10		15		
Val	Ala	Cys	Val	Arg	Ala	Asp	His	Arg	Asp	Arg	Arg	Thr	Ile	Arg	Asp
					20			25				30			

His	Ile	Ala	Met	Ile	Gly	Ile	Leu	Ala	Gly	Ile	Val	Thr	Gly	Pro	Pro	Gly
35																45
Gly	Val	Arg	Gln	Arg	Leu	His	Gly	Ile	Gly	Ile	Ala	Val	Ala	Pro	Gly	
50																60
Pro	Gln	Gly	Val	Val	Vsl	Leu	Ala	His	Gly	Leu	Val	Thr	Gly	Thr	Gly	
65																75
Glu	Val	Gln	Gly	Gly	Arg	His	Val	Ala	Ala	Gly	Val	Val	Asp	Pro		
85																95
Glu	Asn	Gln	Ile	Ile	Arg	Gly	Val	Ile	Gly	Pro	Ala	Pro	His	Asp	Lys	
105																110
Fro	Rsp	Ala	Gly	Ile	Gly	Gln										
																115

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

Arg	Ala	Arg	Gly	Gly	Ile	Arg	Ser	Ser	Lys	Gly	Ser	Arg	Trp	Ser	Arg	Gly	
1																	15
Val	Ile	Gly	Cys	Ile	Ile	Ala	Asp	Ser	Arg	Gln	Ser	Lys	Thr	Ala			
20																	30
Ala	Ser	Fro	Ser	Pro	Ser	Arg	Pro	Gly	Ser	Ser	Ser	Asn	Asn	Ser	Val		
35																	45
Pro	Gly	Ala	Pro	Asn	Arg	Val	Ser	Phe	Ala	Iys	Leu	Arg	Gly	Pro	Leu		
50																	60
Gly	Val	Fro	Gly	Ile	Ile	Asp	Val	Gly	Thr	Asp	Ser	Phe	Gly	Trp	Leu		
65																	75
Ile	Gly	Ser	Pro	Arg	Fcp	Arg	Gly	Gly	Ser	Ala	Ala	Gly	Arg	Gly	Asp	Val	
85																	95
Asn	Fro	Val	Gly	Gly	Leu	Gly	Gly	Val	Ile	Tyr	Gly	Ile	Ser	Pro	Ile		
105																	110
Gly	Asp	Fro	Gly														
																	115

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 811 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(b) TOPOLOGY: [View](#)

(xii) SEQUENCE DESCRIPTION: SEQ ID NO: 189.

TCTGACGCGG CANTCGCTT GGTGAAGATG GTGGATGCG CGCTGCCTGC TGGGGATGCG
 50
 GTGAAAGCC CGGAGGTGT CGCGCGATTC GGCGGAGAACA TGTAACTGCT CAARAGCGTG
 100
 STGCGGGCGG CGGTGATCGC GGYGCGTACG GAGGGCACTG CGACCCACTG TCAACACCAC
 150
 GGGGTTGTC CTTGGCGCTT CGGGTCTTCG TTAAGCTGCTG CGTGACGCCG CGGGCGCGCT
 200
 TCATCGGGTC GGGCGCGATG CGGGCGTACG GGCTGCGGCG TGAAGCGTG GTGGCGCTCG
 250
 AGCGCGTGT CGCGCGCGCG CAGGGCGCAA AGCGCGTGT CGCTGCGCGCG TGGCGAGGGG
 300
 TGTACGTCG CGCGCGCGCG CGCGCGCGCG CGGGCGCGCG CGGTGCGTGC CGGTGCGATG
 350
 ACTTACCGCCCG CATTGGTGGCG GTGGCGCGCG ACGCGCGCGCG CGCACTGCGCG TATGGCGCG
 400
 ACAGCGATT CGCGCGCGCG AGCGCGCGCG CGCGATGCGT CGGGCGCGCG GTGGCGCGCG
 450
 TGGTGTGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGTAACTGCT CGCGCGCGCG
 500
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 550
 GTGGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 600
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 650
 GTGGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 700
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 750
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 800

123 INFORMATION FOR SERV. IN NO. 3021

SECURITY CHARACTERISTICS

- (A) LENGTH: 966 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

```

STCCGGCGT CTGCCCCGAGC ATGACTTTTG ACCAACACCGG CCTCTATGTCG GAGATATCG 60
GACTTTGTGGT TCCCGCTGGC CGGATAGGGC AGCTTGTCGC TTGGCTCGAC GTCACCGTT 120
GCTGGGACUC CGAAACCGTCG CTTCAAGCT AGCTTGCGG TCACACAGT CGGAGGTTA 180
ACGTCACGTT CAATATTCG GTCGATTTC GCCTGAGTC TCTGGCTCGG GACAATCAA 240
GCTTAATCAG TTAGCTGGGAA CGCATTTGGAA CGGGTTGGAT CGGCTTGGGG CTGGTGACCG 300
TGGCGTCCTRA GGTGTCACAGC GCTACCGGCG ACCACGACAT CGGTGTCAC FMTTCACG 360

```

CCAGGACGCGA CGAACGCGC CGGTACAAAG CGGTCTCGGA GGCGTGGGCC GAGGTGGTCG 420
 ACTACCCCGA TCTTGCCTCG CGCTTACGACT CGGGTGACGG CCAATGGTG CGCATCACCG 480
 ACCGAGGACAT CGCGCGCTTG CTGARGAAC CGAGCGCGGA GATCGCGGAG TTGGAGTTTCG 540
 TCGCGCGCGC CGAGTGGAC CGATGAACT TGCGACAGCG CTATTTTG GAGCTGTATT 600
 CGAAGCTTC GAAATGCTAT GTGGCTCTCG CTAAAGACAT CGCGCGACCG CGCGCGATGG 660
 CGATGCTGG TCGCGCGCCG CGCGCGATGG CGAGGAAAT TTAAGCGCCC TATCGACAT 720
 TCGCGCGCGA CGCGCGCGTC CGAAAGACGT AGAGCGATCG AGACGTCG CGATGAACT 780
 CGCGCGCTAT AGAGCGACAT CTGGCGCGTT ATCGAAACGCGA AGATACGCGT AGCTCGACG 840
 CGGGCGCTGC TCGACCGCGAT CGACATTGTT CGCGCGCTCG AGCGCGCTCG CGCGCGCTCG 900
 AAGTGGCGAT CGATCGCGCC CGTCCCCCGA AACCGCGTTC CGGGGGCGGG CGTATCGCGT 960
 CGCGT

(2) INFORMATION FOR EPO AND WIPO:

- (A) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1367 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TECHNOLOGY: linear

ANNUAL REPORT OF THE DIRECTOR OF THE BUREAU OF INVESTIGATION, 1920-1921.

GGGCCACCGCC CGCAATACCC CGAGCGCGACG CCTTACACCCG CTTCGGCGG TTGCCCCCGCT
 60
 TGCGGCGCGT CGCGCGCGGC CGCGCGATG AGTTCTCTTC CGCAAAAGTA CTGGCGTTGC
 120
 CGGGGGAGCC CGCGTTGGGT CGCTTGACCCG CGAGCGCGC GACTCGACG GCGCGAUCCG
 180
 CGTCGCGCGT CGCGCGATG CGCGCGCTGC CGATCGACAT CGCTTGCGG CGACCGTTGC
 240
 CACCCCGGCCG CGCGCGCTGC CGACCGCGCG CGACCGCGAG CGACCGCGG CGCGCGCAC
 300
 CGTCGCGCGT CGCGCGATG CGCGCGCTGC CGACCGCGCG CGACCGCGAG CGACCGCGAC
 360
 TACCCCGACG CGCGCGCGA CGACCGCGAG CGACCGCGCG CGACCGCGAG CGACCGCGAC
 420
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 480
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 540
 CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG
 600
 CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG
 660
 CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG
 720
 CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG
 780

CACGCCGCGT ACCACCCGAA CCACCGCTAC CACGGCTAGA GCTGGAGTC GCGATGTGTA	848
CGAAAGCCCC GGCCTCCGGG CGCCCGCTAC GACCCCGCTG GCGCGCGCT ACACCGCTCG	960
ACCGCTTGCC ACCATCACCAU CGAACGGGGC TCGCATTTTC SCCTCGCGG ACTCCGCGCT	960
GGCGCGCGT GCGCGCGCGG CGACCGGGCG CGCCGGTACG GCGCTGACCA CGCGCGCG	1020
CGGTGGCGT GCGCGGGCT CGCGCGCGG TGGCGACGTC GCGCGCGCGT CGACCGCTCG	1080
GCGTGGCGG AGTCGGATCG CGCGCGCGC CGCGCGCTGA TCGCGATCG	1140
CGGACACATC TCGCGCGCTG TCGCGCGCTC TCGCGCGCG TCGCGCGCTG CGATTGACCG	1200
CGTTTGCGCA CGCGCGCGG CGCGCGCGCG TACCGACCGC CGCCCGATGG CGAACGCGCC	1260
CGGGGTGCG CGCGTGGACG CGCGCGCGCG CGATCGCTGC CGCGCGCGCG CGCGCGCG	1320
CGACGCCGTT CGCGCGCGT CGCGCGCGAC AGCGCGCGT CGACCGCGCA CGCGCGCGCG	1380
CGCGCGCGAC AGCGCGCGCG CGCTTGCCTG CGTCGCGAT CGACCGCGCG CGCGCGCGCG	1440
TGCGCGCGGT TTGACCGCGC CGCGCGCGCG CGCGCGCGCG AGCGCGCGCA TACCGACCG	1500
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG ATCGACCGCG CGCGCGCGCG	1560
CGCCCGAAG CGCGCGCGT CGCGCGATCA CGCGCGCGCG CGACCGCGCG CGAACGCGCG	1620
CTTCAGCTG CGCGCGCGCA CGCGCGCGCG CGTCGCGAC AC TCGCGCGCG CGCGCGCGCG	1680
ACGCTGCCAC CTGCGCGCTG AGCGCGCGAT AC TCGCGCGAC ATCGCGCGCG CGAACGCGCG	1740
CGACCGCGC CGACCGTICA TCGCGCGCG CGACCGCGAC TCGCGCGCG CGCGCGCGCG	1800
CGCGCGCGT AGCGCGCGCTG AGCGCGCGAC CGACCGCGCG CGACCGCGCG CGCGCGCGCG	1860
CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCGCG	1920
CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCGCG	1980
CGCGCGCGCG CGCTCTCTCT GTATCGCGCG CGCGCGCGCG CGCTCTCTCT GTATCGCGCG	2040
CGCGCGCGCG CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCGCG	2100
AGATGATG AGCGCGCGCG AGCGCGCGT GTATCGCGCG CGACCGCGCG CGACCGCGCG CGCGCGCG	2160
CGTCGCGAC CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG	2220
CGCTCGCTTC CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG	2280
CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG	2340
CGCGCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG CGACCGCGCG	2367

(2) INFORMATION FOR SEQ ID NO:202:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 376 amino acids
- (B) TYPE: amino acid
- (C) STRANGENESS:
- (D) TOPOLOGY: linear

(81) SEQUENCE DESCRIPTION: SEQ ID NO:2321

Gin Pro Ala Gly Ala Thr Ile Ala Ala Ser Ser Pro Cys Ala Thr Val
 1 3 10 15
 Gly Ala Gly Gly Gly Thr Gly Ser Pro Val Thr Thr Glu Thr Ala Ala
 20 25 30
 Thr Thr Gly Arg Gly Gly Ser Gly Asp Val Tyr Glu Ser Ala Ala Ser
 35 40 45
 Gly Ala Ala Ala Thr Thr Pro Thr Ala Gly Gly Tyr Thr Val Gly Pro
 50 55 60
 Val Ala Thr Ile Thr Ala Lys Gly Ala Arg Asn Val Ala Leu Arg Asp
 65 70 75 80
 Ser Ala Val Ala Ala Val Ala Ala Ala Thr Gly Ser Gly Gly Thr
 85 90 95
 Ala Val Thr Thr Gly Thr Ala Gly Gly Leu Ala Arg Ala Cys Arg Arg
 100 105 110
 Gly Gly Thr Val Ala Ala Gly Ala Thr Gly Arg Arg Ala Gly Ser Ala
 115 120 125
 Met Ala Ala Arg Ala Ala Val Ala Ala Gly Leu Ile Thr Asp Ala Gly
 130 135 140
 His Ile Cys Arg Ala Val Pro Gly Ala Gly Arg Gly Ala Gly Arg Gly
 145 150 155 160
 Ile Asp Pro Val Cys Pro Gly Glu Ala Gly Ala Ala Gly Thr Thr Gly
 165 170 175
 Ala Ala Met Ala Glu Gln Pro Gly Val Ala Ala Val Thr Ala Arg Thr
 180 185 190
 Pro Asp Ala Cys Gly His Ala Gly Ala Ala Asp Thr Ala Val Ala Ala
 195 200 205
 Val Ala Pro Gln Pro Pro Pro Val Pro Thr Gly Thr Ala Gly Arg Ala
 210 215 220
 Gly Thr Thr Gly Pro Ala Val Ala Ala Val Ala Asp Gln Pro Gly Arg
 225 230 235 240
 Ala Ser Ala Ala Ala Gly Leu Thr Gln Pro Ala Ser Arg Ala Val Ala
 245 250 255
 Thr Val Ala Lys Gln Gln Pro Ala Gly Arg Ala Arg Leu Pro Gly Cys
 260 265 270
 Arg Pro Val Gly Ala Val Ser Asp Gln Arg Ala Pro Gln Lys Arg Leu
 275 280 285
 Gly Gly Arg Ile Ala Arg Thr Gln Gln Thr Pro Leu Asn Ser Gly Phe
 290 295 300

Ber Ala Gly Ile Pro Thr Arg Gly Arg Ser Gln Arg Leu His Lys Leu
 305 310 315 320
 Leu Val Lys Arg Cys His Leu Tyr Ala Glu Arg Leu Ile Leu Pro Ser
 325 330 335
 Met Gly Pro Glu Gln Pro Arg Asn Arg Arg Arg His Phe Ile Gly Ser
 340 345 350
 Arg Ser His His Phe Arg Arg Arg Asp Arg Arg Gly Arg Ile Ser Arg
 355 360 365
 Ala His Leu Arg Thr Asn Ser Arg
 370 375

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2852 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:

GGCCAAAAGC	CCTGGGGCAT	CGGCCCCACG	GAGGCCGCTT	ACGACCAAGAT	GTGGGCCCCAG	60
GGCTGGCGG	CQATETTTGG	CTAACCTTACG	GGGGCTTCGG	GGGCGCTTCG	GGGGTTTGACA	120
CCCTTCGCGC	AGGGGCGGCC	GAACGGTGGG	GGGGCGCGGC	GGCTGGCTCG	GGGGCGCGCG	180
GGCTCAGGTC	GGACGGGGGT	CTTCCGCAC	CTGGGGCTCG	CGAACGTTTG	GGGGGCGAC	240
CTCCCGACG	GTAACTGCGCG	GTAACTTCAT	CTGGGGCTGG	CGAACATCG	CGACGACAC	300
ATGGCGACG	GGACATCGCG	GGCTCTACAC	ATGGGGTTTA	GGAAACGCGG	TCTGGGGTTC	360
AGGGCAGGCC	TGAAACACAG	GGGTTTCCCG	AAACGCGCGA	GGAAACACAT	GGGGTTTGGC	420
AAACGCGCGA	GGAAACACAT	GGGTTTCCGG	AAACGCGCGA	ACCGACACCG	GGTTATGGG	480
CTCAAGGGTA	GGGTTTTCATT	GGGTTTCCCG	GGGCTGAAGT	GGGGCACCGG	GGACGACGTT	540
CTGTTAACCT	GGGGCACCGG	AAACGCGCGA	ATGGCGACG	GGGGTACCGG	GGCTGGGGCG	600
ATGGCGACG	GGGGTACCGG	CTAACACACG	GGTTTTGGCA	ACGGGGCGCA	GGGGAAACACG	660
GGTTTTGGCA	ACGGGGCGCA	CTAACACACG	GGTTTTGGCA	ACGGGGCGCA	CTAACACACG	720
GGTAGCTCGA	ACGGGGCGCA	CTAACACACG	GGTTTTGGCA	ACGGGGCGCA	CTAACACACG	780
GGCTACCTGA	ACGGGGCGCA	CTAACACACG	GGTTTTGGCA	ACGGGGCGCA	TGTGACACCG	840
GGGCGCTTCA	TTACTGGCAA	CTAACACACG	GGCTTTCTGT	GGGGGGGUGA	GGGGAAAGGC	900
CTGATTTGCG	GGGGGGGGCG	CTTCCTTCACG	TGAGGAGTGT	GGGGGGGGCA	GGGGATCTTC	960

AACAGGGGTT CGCGTTCGCG CTCGAACTCGG CTGGCAACAA TTCTGGCTC
 1029
 TTCAACTCTT CGTCGGGGGC CATGGTAAC TCCGGGCTGG CAACGGGGG CCTGCTGGTA
 1030
 TCGGGCTGTA TCACCTCGGG CAACACCGTA TCGTTTGTGTC TCAACATGAC CGCTGGTGGC
 1140
 ATCACACACG CGGGCTTGAT CTGGGGCTTC TTCAACACCGA AAAGGACATC GTGGGATTC
 1280
 TTGGTGGGC CACGGCTCTT CAATCTGGC CTGGCAACAC GGGGGTGTG GAACATTTC
 1281
 GGCACACCA ACATGGGAA TTACACATT CGGGCAAGG CGACGCTGG TGACTTCAA
 1328
 ATCCCTGGCA CGGCAACAT CGGGGGCAA AACATCTTG CGACGGCA CGCTGGGAGC
 1389
 TTCAATATCG CGAGTGGAA CATGGGATA TTCAATATCG GTTCCGGGAAG CCTGGGURAC
 1440
 TACACATCG GTTGGGAA CGTGGGGATC TACACATCG GTTGGGAAAG CGTGGGAC
 1500
 TACACCTCG GTTGGGAA CGTGGGGAA TACACATCG CGACGGGTT GTGGGGGAAAC
 1560
 AACCAACGCG CGTGGGAAAC CGGGGGAA TACACATCG CGACGGGTT GTGGGGGAAAC
 1620
 AACCAACGCG CGTGGGAAAC CGGGGGAA TACACATCG CGACGGGTT GTGGGGGAAAC
 1680
 TTCAATTGCG CGACGGGAAAT CGTGGGAAAT TTCAACGGCG CGACGGGAAAT CGTGGGAAAT
 1740
 CGACGGGAAAT CGTGGGAAAT CGGACGGGAAAT CGACGGGAAAT CGTGGGAAAT
 1800
 TTCAATCGT CGAGCTCAA CGGGGGATC CTCAACGGCG CGAGCTTCA CGGGGGCTTC
 1860
 TACACACCG CGACGGGAAAT CGGGGGAAAT TTCAACGGCG CGACGGGAAAT CGGGGGAAAC
 1920
 TTCAACGGCG CGACGGGAAAT TACACACCG CGACGGGAAAT CGGGGGAAAC
 1980
 TTCAATCGT CGACGGGAAAT CGGGGGAAAT TTCAACGGCG CGACGGGAAAT CGGGGGCTTC
 2040
 TTGGTGGCG CGGATAACCA CGGGGGAAAT CGACGGGAAAT CGTGGGAAAT CGTGGGAAAT
 2100
 ATGGCTTAAAC AGGGGGAAAT CGTGGGAAAT CGACGGGAAAT CGGGGGAAAC
 2160
 ATGATLACKG CGACGGGAAAT CGTGGGAAAT TTGGGGAAAT CGGGGGAAAT CGGGGGAAAT
 2220
 TTCTTCTTGG CGGGGGAAAT CGTGGGAAAT TTGGGGAAAT CGGGGGAAAT CGGGGGAAAT
 2280
 ACCATGGCG CGGGGGAAAT CGGGGGAAAT ATGAGGATTC CGGGGGAAAT CGGGGGAAAC
 2340
 ACCATTAACG CGGGGGAAAT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2400
 TGGTGGCGT TTTCACCTCG GGGCACCGGT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2460
 CGGGGGAAAT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2520
 AACCTGGGT CGGGGGAAAT CGGGGGAAAT AACCTGGGT CGGGGGAAAT CGGGGGAAAC
 2580
 AACACGGCTA CGGGGGAAAT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2640
 AACGGGGCTT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2700
 AACCTGGGT CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2760
 CGGGGGAAAT CGGGGGAAAT CGGGGGAAAC
 2820

STARBUCK MANGAINTS CONFIDENTIAL

2642

(2) INFORMATION FOR SDO IS REV-206.

5.2. SHOCKING CHARACTERISTICS.

- (A) LENGTH: 943 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(81) SEQUENCE DESCRIPTION: SEQ ID NO: 204;

Gly Glu Asn Ala Pro Ala Ile Ala Ala Thr Glu Ala Ala Tyr Arg Gln
35 36 37 38 39 40 41 42 43 44 45

Met Trp Ala Cys Asp Val Ala Ala Met Phe Gly Tyr His Ala Gly Ala
29 29 30

Ser Ala Ala Val Ser Ala Leu Thr Pro Phe Gly Gln Ala Leu Pro Thr
35 40 45

Val Ala Gly Gly Gly Ala Leu Val Ser Ala Ala Ala Ala Glu Val Thr
50 55 60

Thr Arg Val Phe Arg Asn Ileu Gly Ileu Ala Asn Val Arg Glu Gly Asn
65 70 75 80

Vai Arg Asn Gly Asn Val Arg Asn Thr Asn Leu Gly Ser Ala Asn Ile
85 90 95

Gly Asn Gly Asn Ile Gly Ser Gly Asn Ile Gly Ser Ser Asn Ile Gly
100 105 110

Phe Gly Asn Val Gly Pro Gly Leu Thr Ala Ala Leu Asn Asn Ile Gly
115 120 125

Phe Gly Lys Thr Gly Ser Asn Asn Ile Gly Phe Gly Asn Thr Gly Ser
130 134 140

Asn Asn Ile Gly Phe Gly Asn Thr Gly Asp Gly Asn Arg Gly Ile Gly
145 150 155 160

Leu Thr Gly Ser Gly Leu Leu Gly Phe Gly Gly Leu Asn Ser Gly Thr
165 170 175

Gly Asn Ile Gly Leu Phe Asn Ser Gly Thr Gly Asn Val Gly Ile Gly
180 185 190

Asn Ser Gly Thr Gly Asn Thr Gly Ile Gly Asn Ser Gly Asn Ser Tyr
195 200 205

Asn Thr Gly Phe Gly Asn Ser Gly Asp Ala Asn Thr Gly Phe Phe Asn
219 215 220

Ser Gly Ile Ala Asn Thr Gly Val Gly Asn Ala Gly Asn Tyr Asn Thr
225 230 235 240

Gly Ser Tyr Asn Pro Gly Asn Ser Asn Thr Gly Gly Phe Asn Met Gly

285	290	295
Gln Tyr Asn Thr Gly Tyr Leu Asn Ser Gly Asn Tyr Asn Thr Gly Leu 260 265 270		
Ala Asn Ser Gly Asn Val Asn Thr Gly Ala Phe Ile Thr Gly Asn Phe 275 280 285		
Asn Asn Gly Phe Leu Trp Arg Gly Asp His Gln Gly Leu Ile Phe Gly 290 295 300		
Ser Pro Gly Phe Phe Asn Ser Thr Ser Ala Pro Ser Ser Gly Phe Phe 305 310 315 320		
Asn Ser Gly Ala Gly Ser Ala Ser Gly Phe Leu Asn Ser Gly Ala Asn 325 330 335		
Asn Ser Gly Phe Phe Asn Ser Ser Gly Ala Ile Gly Asn Ser Gly 340 345 350 355		
Leu Ala Asn Ala Gly Val Leu Val Ser Gly Val Ile Asn Ser Gly Asn 355 360 365		
Thr Val Ser Gly Leu Phe Asn Met Ser Leu Val Ala Ile Thr Thr Pro 370 375 380		
Ala Leu Ile Ser Gly Phe Phe Asn Thr Gly Ser Asn Met Ser Gly Phe 385 390 395 400		
Phe Gly Gly Pro Pro Val Phe Asn Leu Gly Leu Ala Asn Arg Gly Val 405 410 415		
Val Asn Ile Leu Gly Asn Ala Asn Ile Gly Asn Tyr Asn Ile Leu Gly 420 425 430		
Ser Gly Asn Val Gly Asp Phe Asn Ile Leu Gly Ser Gly Asn Leu Gly 435 440 445		
Ser Gln Asn Ile Leu Gly Ser Gly Asn Val Gly Ser Phe Asn Ile Gly 450 455 460		
Ser Gly Asn Ile Gly Val Phe Asn Val Gly Ser Gly Asn Leu Gly Asn 465 470 475 480		
Tyr Asn Ile Gly Ser Gly Asn Leu Gly Ile Tyr Asn Ile Gly Phe Gly 485 490 495		
Asn Val Gly Asp Tyr Asn Val Gly Phe Gly Asn Ala Gly Asp Pro Asn 500 505 510		
Gln Gly Phe Ala Asn Thr Gly Asn Asn Asn Ile Gly Phe Ala Asn Thr 515 520 525		
Gly Asn Asn Asn Ile Gly Ile Gly Leu Ser Gly Asp Asn Gln Gln Gly 530 535 540 545		
Phe Asn Ile Ala Ser Gly Trp Asn Ser Gly Thr Gly Asn Ser Gly Leu 550 555 560		
Phe Asn Ser Gly Thr Asn Asn Val Gly Ile Phe Asn Ala Gly Thr Gly 565 570 575		

Asn Val Gly Ile Ala Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn
 580 585 590
 Pro Gly Thr Asp Asn Thr Gly Ile Leu Asn Ala Gly Ser Tyr Asn Thr
 595 600 605
 Gly Ile Leu Asn Ala Gly Asp Phe Asn Thr Gly Phe Tyr Asn Thr Gly
 610 615 620
 Ser Tyr Asn Thr Gly Gly Phe Asn Val Gly Asn Thr Asn Thr Gly Asn
 625 630 635 640
 Phe Asn Val Gly Asp Thr Asn Thr Gly Ser Tyr Asn Pro Gly Asp Thr
 645 650 655
 Asn Thr Gly Phe Asn Pro Gly Asn Val Asn Thr Gly Ala Phe Asp
 660 665 670
 Thr Gly Asp Phe Asn Asn Gly Phe Ile Val Ala Gly Asp Asn Gln Gly
 675 680 685
 Gin Ile Asn Ile Asp Leu Ser Val Thr Thr Pro Phe Ile Pro Ile Asn
 690 695 700
 Glu Gin Met Val Ile Asp Val His Asn Val Met Thr Phe Gly Gly Asn
 705 710 715 720
 Met Ile Thr Val Thr Gln Ala Ser Thr Val Phe Pro Gln Thr Phe Tyr
 725 730 735
 Leu Ser Gly Leu Phe Phe Gly Pro Val Asn Leu Ser Ala Ser Thr
 740 745 750
 Leu Thr Val Pro Thr Ile Thr Leu Thr Ile Gly Gly Pro Thr Val Thr
 755 760 765
 Val Pro Ile Ser Ile Val Gly Asn Leu Gln Ser Arg Thr Ile Thr Phe
 770 775 780
 Leu Lys Ile Asp Pro Asn Pro Gly Ile Gly Asn Ser Thr Thr Asn Pro
 785 790 795 800
 Ser Ser Gly Phe Phe Asn Ser Gly Thr Gly Gly Thr Ser Gly Phe Gln
 805 810 815
 Asn Val Gly Gly Ser Ser Gly Val Thr Asn Ser Gly Leu Ser Ser
 820 825 830
 Ala Ile Gly Asn Ser Gly Phe Gln Asn Leu Gly Ser Leu Gln Ser Gly
 835 840 845
 Trp Ala Asn Ile Gly Asn Ser Val Ser Gly Phe Phe Asn Thr Ser Thr
 850 855 860
 Val Asn Leu Ser Thr Pro Ala Asn Val Ser Gly Leu Asn Asn Ile Gly
 865 870 875 880
 Thr Asn Leu Ser Gly Val Phe Arg Gly Pro Thr Gly THR Ile Phe Asn
 885 890 895
 Ala Gly Leu Ala Asn Ile Gly Gln Leu Asn Ile Gly Ser Ala Ser Cys
 900 905 910

Arg Ile Arg His Glu Leu Asp Thr Val Ser Thr Ile Ile Ser Ala Phe
 915 928 929

Cys Gly Ser Ala Ser Asp Glu Ser Asp Pro Gly Ser Val Ser Cys
 930 935 940

(2) INFORMATION FOR SEQ ID NO:205:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GGATCCGAT GGCCTCATCAT CATCATCATC ACCTGATGGA CATCATCGG ACC 53

(2) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:206:

OCTGAAATTCA GGCCTCGGT GCGCCGCGCT CATCTTGAC GA 42

(2) INFORMATION FOR SEQ ID NO:207:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:207:

GGATCTGCA GGCCTGAGG CACCGAGCC T 31

(2) INFORMATION FOR SEQ ID NO:208:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:208:

CTCTGAAATTG AGGGCTGGAA ATCGCTGGGA T

31

(2) INFORMATION FOR SEQ ID NO:209:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:209:

GGATCYACGCG CTGAGATGAA GACCCATGCC CCT

33

(2) INFORMATION FOR SEQ ID NO:210:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:210:

GGATACTCGC AGRATTTCGG TTTCRAAGCCC RTTTCUGA

38

(2) INFORMATION FOR SEQ ID NO:211:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xiv) SEQUENCE DESCRIPTION: SEQ ID NO:211:

CCGATGCAA GCGCACTTCCC CACAAAGGGGC

39

(2) INFORMATION FOR SEQ ID NO:212:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

CTTGATGGAA TTGTGAGGCC GGTAAAGGTCG GCTGGGGT

37

(2) INFORMATION FOR SEQ ID NO:213:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1676 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:213:

TCGGAAATGG GACGGCGCCG CTAGCGGCCG ATTAAACGCC CGGGGTGTTGG TGGTATGCCG	60
CAGCGTGACC GCTACACTTG CCAAGGCCCT AGCGCGCGCT CCTTTGGTT TCTTCGCTTC	120
CTTTCTGCCG ACGTTGCCCG GCTTTCGCCG TCAAGCTCTA AATGGGGGC TCCGTTTAGG	180
GTTCCGATT AGTGGTTAAC GGAACCTCGA CCCCCAAAAA CTGGATTTGG STGATGGTTG	240
ACGTAATGGG CCTACGCCCT CTAGACGGT TTTCGCGCT TTGACGTTG AGTCCACGGT	300
CTTTAATAGT CGACTCTPTG TCCBAACTCG AGAAACACTC AACCTATCT CGTTCTATTG	360
TTTGATTTA TAAGGGATTG TGCGGATTC GGCGTATGCG TTAAGGATTC AGCTGGTTA	420
ACA AAAAATTG AAAGGAAATT TTAACAAAAT ATTAAACCTT ACATTTGAG GGGGCACTT	480
TGGGGAAAT GTGGGGGAA CCCTATTTG TTFAATTITTC TAATGACCTT CAATATGTA	540
TCCGCTCTG AATTATCT TAGAAAAACI CATGGCGAT CAATGAAAC TGGATTTAT	600
TCAATGCG ATTATAATAA GGTATTTTG GAAACCCG TTTCGTTAAAT GAAAGGAAAT	660
ACTCACCGAG GCGTTCTAC AGGATGGCAA GATCCTGTA TCGGCTGCG ATTCGGACTC	720
CTCCGACRTC AATACAGCTT ATAATTTCCG CCTGTCGAA AATAGGCTA TCAATGAGA	780
AATGCGATC ACTGAGCAT GAACTCCGTA AGCATGGCGA AAGTTTGGG ATTTCGTTCC	840
AGACTCTTC AAGCGGCGAG CCTTATCTCTG CCTCATCAAA ATCACTCGA TCAACCAAC	900
CCTTATTCAT TGTGTGTTGC GCGTGGCGA GCGGAAATAC GCGATCGTG TTAAGGAGC	960
AATTACAAAC AGGATGGCGA TCGAACCGGG GCGGCGACG TCGGCGCGA TCAACATAT	1020
TTTCACCTGA ATCAGGATAT TCTTCTAATA CCTGGCATGC TGTGTTCCCG GGGATCGCG	1080
TGGTGAGTAA CCTGCGATCA TCAGGAGTAC GGTAAATAG CTTGATGTCG CGGAGGCGA	1140

TAAATTCCGT CAGCAGCTT AGTCGACCA TCTCATCTGT AACATCATTC GCAAGCGTAC 1200
 CTTCGGCTA TTTCAGAACG AACTCTGGCG CATCGGGCTT CCCTACATT CGATAGCTT 1240
 TGCGACCTGA TTGCCCGACA TTATCGCGAC CCGATTTATA CCCTATTTAA TGCGATCTA 1280
 TCTTGCGATT TAATTCGCGCCT CTAGAACAAAG AGCTTTCGCG TTGAATATGC CTCATAAACAC 1320
 CCCTGTATT ACTTGTTATG TAAGCAGACG GTTTTATGTG TCAATACCAA ATCTGTTAA 1360
 CGTGRAGTTTT CGTTCGACTG AGCTTGCGAC CGCTTAAGAA AGCTTAAGG ATCTTCTTGA 1400
 GATCCTTTT TTCTCGCGCT AATCTGCTC TTGCAAAACAA AAAAACGCGC CGTACCGCG 1440
 GTCGGTTGTTT TGGCGGATCA AGAGCTACCA ATCTTTTTC CGAGGTAAAC TGGCTTGGC 1480
 AGAGCGACGA TACCAAAAC TGCTCTTCTA CGTACGGCGT ACCTTGGCGA CGCTTCGAG 1520
 AACCTGTTAG CAIXXCGTAC ATCTCGCGT CGCTTATCC TGTAACTAGT CGCTTCGCG 1560
 AGTGGCGATA AGCTGTTCTC TACCGGGTTG GATCTAACAC GATACTTACG GGATAAGCG 1600
 CAGCGGTGCG GTCGAACCGG CGTTCGCTGC AGACAGGCCG CGTGGAGCG AACGACCTAC 1640
 ACCGAACTGA GATACTTACA CGCTGAGCTA TGAGAAGCG CGACGCTTC CGAGGGGAGA 1680
 AAGGGGGAGA GGTATCGCTG AGCGCGCAOG CTGCGAACAG CGAGCGCGCG CAGGGAGCTT 1720
 CCAGGGGGAA AGCGCTGCA TGTATTAATG CGCTCGCGT TTGCGACCT CGTCTTGAG 1760
 CGTCAATTCTT TGTGTCGTC CGTACGGGGCG CGAGCTTATG CGAAAACCG CGCGAACGGCG 1800
 CGTCTTTCAG CGTTCGCTGC CTTPCTGCTG CGTCTTTCGTC ACATTTCTT TCGTGCCTTA 1840
 TCGCTGATT CTGIGKATAA CGCTTACCG CGTCTTCACT CGCTTCGATC CGCTCGCGC 1880
 AGCGAACGAC CGCGCGCGAG CGAGCTTACG AGCGAGGAGA CGAGAGCGCG CGCTGATCGG 1920
 TATTTCTCG TTAACGATCTG CGGGGGTATT TCACAGCGCA TATTTGGCG ACCTCTGAGTA 1960
 CGATCTGCTC TGATCGCGCA TGTAAAGC AGTATACACT CGCTCTACCG TACGTGACTG 2000
 CGTCAGGGCT CGCCCGCGAC AGCGCGCGAC AGCGCTTACG CGCGCTGCG CGGCTTCTCT 2040
 CGTCGGCGCA CGCGCTTACG CGAGCTTACG CGAGCTTACG CGAGCTGCG CGTGTGCG 2080
 CGTCTTCCAG CGTACGCGCA AGCGCGCGAG CGAGCTCGCG TAAAGCTCAT CGCGCTGCG 2120
 CGTGAAGGGAT CGACGAGATG CTGGCTGTTG ATCCGGCTGC AGCTGTTGA GTTCTCGAG 2160
 AGCGCTTATG CGTCTGCTTC TGAATACCG CGCTTACGTTT CGCGCTGTTT TTGCGCTTCT 2200
 CGCTACTATG CGCTCCGCTG AGCGCGCGAT CGTGTGCTG CGGCTAATGA TACCGATGAA 2240
 AGCGAGAGCG ATGCTTACGCA TACGGCTTAC TGATGAGAA CGAGCGCGT TACTGAGCG 2280
 TTGTAGCTT ATACAGACTGC CGTATGGAT CGCGCGCGAC CGAGAGCGAA TGCGTACGG 2320
 TCAATCGCG AGCGCTGCTTA ATACAGATG AGGTGTTGCA CGCGCTACCG AGCGCTGCTC 2360
 CGCGCTGCG AGCGCTGACCA TAATGTTCA CGCGCTGAGC TTCCCGTTT CGAGCTTTA 3000

CGAAACAGGG AAACCGAAGA CCTTCATGT TTTTGTCAGG CTGGCAGAGC TTTCGACCA
 GCAGTGGCTT CAGGTGGCT GGCCTATGGG TTTATTTCATC TGCCTAACCA TTAAGCCAC
 CGGCCAGCTT AGCCGGTCC TCAAGGCGAA TGGCACGATC ATGGCGCCCT GTGGGGGCTGC
 CATCCGGCGG ATATTCGGCT GCTTCCTCCC GAAAGGTTT GTGGGGGCGA TACTKACCAA
 GGTTTGGAGG AAGGGCTGCGA AGATTCGGA TGGCGAACG GAGCGGCGA TCACTTGCG
 GCTTCAGKIA AGGGCTTCTT CGGGAAAT GACCGAGG GCTGGCGCGA CGCTGCTTAC
 GAGTGGCAT ATARAGAAGA CAGTCATAAG TGGCGGAGG ATAGTCATCG CCCTGGCGCA
 CGGGAGGAG CTGACTTGGT TGAGGGCTT GAGGGCATC GTCGAGAGTC CGGGCGCTA
 ATGGTGGACG TAACTTACCT TTTTGGTTT CGCTCTTCG CGGGCTTCG AGTCGGGAA
 CTGTCGTCG CAGCTGCTT ATTAACTTG CGACGGGGG GGGAGGCGG GTTGGCGTAT
 TGGGGGGGGG GGTGGTTTCTT CTTTCACCA GTGAGGAGGG GAAAGCTGA TGGCGCTTC
 CGGGCTGGCC CGAGAGAGT TGGAGGAGG GGTGGCGCTT GGTGGCGCC AGCGGGCGAA
 AACCTCTTT GATGGTGGT AACGGGGGG TATACATCA GCTGGCTCTG CTATCTCTGT
 ATCCACTAC CGAGATATCC GCAACGACG GCAAGGGCGA CGGGTGTATG CGCGCGATC
 CGCCCAAGCG CAGCTGCTTGG TGGCGAACG GCAAGGAGCT CGGGAGCGATG CGCTCTTC
 GCAATTGCA GGTGGTTGAA AACGGGAGA TGGACTCA GTCGGCTTCG CGGGCGCTA
 TGGKKTGAT TTGATGGCGA GTGAGMTAT TATGCCAGG AGCGGAGGG AGAGGGCGCG
 AGCGAGAGT TAATGGGGCG GCAACGACG CGGGCTTCG GTCAGGGCAAT CGGGCGCTAT
 GGTCCACCGG CGGGCGCTA CGGGCGCTT GGGGAAATG ATAGCTTG ATGGCTCT
 GGTGGAGAG ATCAAGAAAT AACGGGGAGA CATTAGCTCA GCAAGGCTTC ACACCAATGA
 CATGCTGTC ATCGAGGGGG TAGTTAATGA TCAAGGCGCT GGGGGGTTGC CGGGGAGAT
 TGTGGACCC CGGGTTCAGG GGTGGAGCG CGCTCTTCG TACCTCGAG ATGACCGCC
 TGGCGGGGG TTGATGGCG GGGGATTA TGGCGGGAG AACGGGGAGC GGGGGCGCGA
 GGGGGAGCT GGGGGGGGG AGCGGGAGCA GCAAGGAGTC TTTGGGGCGG AGTTGTTGCG
 CGACCGGGTT GGGGATTA TTGAGGCTG CGAGGGGGG TTGGACTTTT TGGCGCTT
 TGGAGAGAG ATGGTGGCT TGGTGGAGA CGGGGGAGA GGTGGCGATG GGGGGCGCG
 CAGCTGTCG GACGGCTAT AACGGCTTCG GTTGCACATT CGGGCGCTG ATGGCTCT
 CTGGGGGG CGTGGATGCG ATACGGCGAA AGGTTTTTGGC GTTGGGGCGG GTGTCGGCGA
 TGGAGAGCT GGTGGCTTGG CGGGCGCTTCG ATGGAGGAGG AGGGGAGCTAG TGGGGCG
 CGGGGGAGCA CGGGGGGGCG AGGGGGAGCT GCAAGGAGG AGGGGGGGCG GAGGGGGCG
 CGGGGGAGCG GGTGGGGCGAG CGTACGGAGG AGGGGGGGCG GAGGGGGCG GAGGGGGCG

CGAGCCCCAT	CATCCOCATC	GCTGTTTGTG	GCGATATAAG	CCTCCACCAAC	CGCACCTGTC	4929
GGCCGCGTGA	TGCUGGCCAC	GATGCGTCCG	GCGTTCAGGA	TCCAGATTCG	GATCCCGCGA	4930
AAATTAATACG	ACTUACATATA	GGGGAAATTGT	GAGCCGATTA	CAATTCCCT	CTAGAAATTA	5040
TTTGTTTAA	CITTAAGAAC	GAGATATACA	TATGGGGCAT	CAUCACATC	ACACAGTGTAT	5190
CGRCRCAAC	GGGACGAGCC	CCACATCTTG	GGAAACACCG	GGGGCGGAGG	GGGTCGAGCG	5169
GGGGGGGGAT	AGCGTCGATG	ACATCCGGGT	GGCTCGGGTC	ATGGAGCGGG	ACATGGCCGT	5226
GGACRGCGCC	GGCAAGAGCA	CTTAXLQAT	CAAGCTGCAA	GTGTCGTTCA	AGATGAGGCC	5280
GGCCGAAACG	AGGGGCTGAA	AACUUCCGAG	CTTCGCGGT	GAACGGGGGT	CGGGGGCGG	5345
TACTTCGGG	ACTTCGCGG	CTTCGTCGCG	GGTGAACGTT	GGCGAGACCG	GTACACCCGT	5468
GTCTCACCOS	CTGTCFAACC	TGTCGGGGCG	GGCTTTTCAC	GGAGGTTTC	CGAACGTCAC	5469
GTCAACCGT	CAGGGGGCG	GTTCGTCGTC	GGGGATGCGG	CGAGGGGGCG	CGGGGGCGGT	5539
CAACATTCGG	GGCTCCCGCG	CTTATCTTCG	GGAAAGGGGT	ATGGGGCGGC	ACAAGGGGT	5580
GATGAACTC	GGGTAGCGA	TCTCCGCTCA	GGAGCTCAG	TACAACCTGC	CCGGAGTGAAG	5640
CGAGCACCCT	AAGCTGAAAC	GAAGGGGGT	GGGGGGCTG	TACGGGGGA	CCATCAAAAC	5700
CTGGGACGAC	GGCCACATCG	CTGCGCTAA	CCGGGGGCG	AACTGGGG	CGACCGCGGT	5768
TTTGTGCGT	CAACCTCGCG	AGGGTTCGG	TAACACCTTC	TGGTTCACCG	AGTACGTTTC	5828
CAGCAAGAT	CCCAGGGGCT	GGGGGAACTG	GGCGGGGTTG	GGCGACCGG	TGAGCTTCG	5888
GGCGTTCGG	GGTGGGCGGTC	CTGAGACCGG	CAACGGGCG	ATGGTGCAGG	GGTTCGGCGA	5940
GGACCGGGCG	TGCGTGGGCT	ATATCGCCAT	CACTTCCTTC	GGCGGGGCGA	CTTCAGGGCG	6000
ACTCGGCTAG	GGCCACACAG	GGTATGGCT	TGGGTTTC	TGGTGGCGCG	ADGGGCGRAAG	6060
CATTCAAGGC	GGGGGGGCTG	GTTCGTCGTC	GAAGACCGG	GGAAACCCAG	CGATTCGGAT	6120
GATCGACGG	GGGGGGGGCG	AGGGTCAAC	GATCACTAAC	TACGGTACCG	CCATGGCGAA	6180
CTACCGGGCA	GGGGGGGGCG	CCACCGCGCA	GGCTTTCAG	GGTTTTCTGC	ACTGGGGCGT	6240
CGCGACGGCG	AAACAGGGCT	CTTCTCTCGA	GGAGGTTCT	TTCGAGACCG	TGCGGGGGCG	6300
GGTGTTGAG	TTTGTGCGT	GGTGTGCGT	GAAGTTTGC	ATGGGTTGAGA	TGAGTCACCG	6360
TGCGCTTAC	CTGGGACGAG	AGGGGCTAA	TTCGAGGGG	ATTCGGGGCG	ACCTGGACAC	6420
CCAGATCGAC	CAAGGGGAGG	TTGGTTGCGA	GGCGAGGGCG	GGGGGGGGCG	GGGGGGGGCG	6480
GGGGAGGGCG	GGCCACACCGG	GGGTGTTGCG	CTTCCGAGAA	GGAGGCGAT	ATGGGAGGCA	6540
GGAACTCGAC	GAATCTTGA	GGATATTCG	TGCGGGGGCG	CTTCGATACG	GGGGGGGGCG	6600
GGAGGAGCG	GGGGGGGGCG	TCTCCGCGCA	ATGGGGCTTG	GTGGGCGACG	GGGGGGGGCG	6660
GGGGGGGGCG	ACGGCTCGAC	GGGGGGGGCG	ADGGGCGCG	CTTGTGCGCG	GGGGGGGGCG	6720

GGCGCCGCGCG AACACAGCGGA ATGCCCGAGG GGGGGATCCC AACCGGAAACG CTGCGCGCGG 6780
CGACCGGAAAC GCACCGGCGC CACUTGTCTA TCCCCAAAC GCACCGAACG CTGTCGCGAT 6840
GCACAAACCG GTTGGGAGGT TCACTTCCG GCTGGCTGTG CGCTGGGTTGG AGCTTGACGC 6900
CGCGCATTC GATCTACGGT CACGCTCGT CGCGAAACG AGCGGAGACG CGCGATTTGC 6960
CGGACAGCGC CGGGCGGTTGG CAAATGACG CCGTATCGTG CTGGCGCGCG TAGACCAAAA 7020
GCTTACGCC AGCGCGGAGAC CGACCGACTG CGACCGCGCG CGCGGTTGG CGCTGGGACAT 7080
GGTAAATTC TATATGCCG AGCGCGGACG CGCGAAACG CGGGAAACCG TCTCGGTTGA 7140
CGCGCAACGGG GTGCTCTGGAA CGCGCGYCGTA TCACTGACTG AATTCGAGG ATCGCGAATA 7200
CGCGAAACGGG TCAATCTGCA CGGGCGTAAT CGCGCTCGGCC CGCGCGGAGG CACCGGACCG 7260
CGGGCGCGCGT CACGCTCGTG GTTGGGATGG CGCGACAAACG CGCTGGGACAA 7320
GGGGCGCGCC AGCGCGCTGG CGCGATCGT CGCGGTTGG ATCGCGCGCG CGCGCGCGCGCG 7380
GGCGCGCGCT CGTGCAGAGC CGCTCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 7440
CGCGCGCGCA CGCACCGCGC AGCGGACCGT AGCGGCGCTGA GATCTTGCGA GATATCGATC 7500
ACACTGGCGC CGCTCGAGG AACACCGCGA CGACCGCTGA GATCGCGCGT CTACAAAGG 7560
CGCGAAAGGA CGCTCGATGG CGTGCAGCGC CGCGGAGGAA TACGCTCGATG AACCGCTGG 7620
GGCTCTGAA CGCTCGATGG CGCGGTTGGT GTGCGAAAGGA CGCGATCGATG CGCGAT 7680

(2) INFORMATION FOR SED ID NO: 216:

(3) SUPPORTIVE CHARACTERISTICS:

- (A) LENGTH: 802 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

and 3. Sequence differences seen in 16S rRNA

```

Met Gly His His His His His Val Ile Asp Ile Ile Gly Thr Ser
          5                   10                  15

Pro Thr Ser Trp Glu Gln Ala Ala Glu Ala Val Gln Arg Ala Arg
          20                  25                  30

Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Glu Gln Asp Met
          35                  40                  45

Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Gln Val
          50                  55                  60

Ser Phe Lys Met Arg Pro Ala Gln Pro Arg Gly Ser Lys Pro Pro Ser
          65                  70                  75                  80

```

Gly Ser Pro Gln Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro
 85 90 95
 Ala Ser Ser Pro Val Thr Leu Ala Gln Val Gly Ser Thr Leu Leu Tyr
 100 105 110
 Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Gln Arg Tyr Pro Asn
 115 120 125
 Val Thr Ile Thr Ala Gln Gln Thr Gly Ser Gly Ala Gly Ile Ala Gln
 130 135 140
 Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser
 145 150 155 160
 Glu Gly Asp Met Ala Ala His Lys Leu Met Asn Ile Ala Leu Ala
 165 170 175
 Ile Ser Ala Gln Gln Val Asn Tyr Asp Leu Pro Gln Val Ser Gln His
 180 185 190
 Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile
 195 200 205
 Lys Thr Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gln Val Asn
 210 215 220
 Leu Pro Gly Thr Ala Val Pro Leu His Arg Ser Asp Gly Ser Gly
 225 230 235 240
 Asp Thr Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Gln Gly
 245 250 255
 Trp Gly Lys Ser Pro Gln Phe Gly Thr Thr Val Asp Phe Pro Ala Val
 260 265 270
 Pro Gly Ala Leu Gln Gln Asn Gln Gly Met Val Thr Gly Cys
 275 280 285
 Ala Gln Thr Pro Gln Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp
 290 295 300
 Gln Ala Ser Gln Arg Gly Leu Gly Gln Ala Gln Leu Gln Asn Ser Ser
 305 310 315 320
 Gly Asn Lys Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala
 325 330 335 340
 Gly Phe Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp
 340 345 350
 Gly Phe Ala Pro Asp Gln Tyr Pro Ile Ile Asn Tyr Gln Tyr Ala Ile
 355 360 365
 Val Asn Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala
 370 375 380
 Phe Leu His Trp Ala Ile Thr Asp Gln Asn Lys Ala Ser Phe Leu Asp
 385 390 395 400
 Gln Val His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp
 405 410 415

Ala Leu Ile Ala Thr Ile Ser Ser Ala Glu Met Lys Thr Asp Ala Ala
 420 425 430
 Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile Ser Gly Asp Leu
 435 440 445
 Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly Ser Leu Gln Gly
 450 455 460
 Gin Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala Ala Val Val Arg
 465 470 475 480
 Phe Gin Glu Ala Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu Ile Ser
 485 490 495
 Thr Asn Ile Arg Gln Ala Gly Val Glu Tyr Ser Arg Ala Asp Glu Glu
 500 505 510
 Gin Gln Gln Ala Leu Ser Ser Gin Met Gly Phe Val Pro Thr Thr Ala
 515 520 525
 Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro
 530 535 540
 Val Ala Pro Pro Pro Pro Ala Ala Asn Thr Pro Asn Ala Gin Pro
 545 550 555 560
 Gly Asp Pro Asn Ala Ala Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro
 565 570 575
 Pro Pro Val Ile Ala Pro Asn Ala Pro Gin Pro Val Arg Ile Asp Asn
 580 585 590
 Pro Val Gly Phe Ser Thr Ala Leu Pro Ala Gly Trp Val Glu Ser
 595 600 605
 Arg Ala His His Phe Asp Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr
 610 615 620
 Gly Asp Pro Pro Phe Pro Gly Glu Pro Pro Pro Val Ala Asn Asp Thr
 625 630 635 640
 Arg Ile Val Leu Gly Arg Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu
 645 650 655
 Ala Thr Asp Ser Lys Ala Ala Arg Ser Leu Gly Ser Asp Met Gly Gln
 660 665 670
 Phe Tyr Met Phe Tyr Pro Gly Thr Ala Ile Asn Gln Glu Thr Val Ser
 675 680 685
 Leu Asp Ala Asp Gly Val Ser Gly Ser Ala Ser Tyr Tyr Gln Val Lys
 690 695 700
 Phe Ser Asp Pro Ser Lys Pro Asn Gly Gln Ile Trp Thr Gly Val Ile
 705 710 715 720
 Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp
 725 730 735
 Phe Val Val Trp Leu Gly Thr Ala Asn Asn Pro Val Asp Lys Gly Ala

211

746

745

750

Ala Lys Ala Leu Ala Glu Ser Ile Arg Pro Leu Val Ala Pro Pro Pro
755 760 765

Ala Pro Ala Pro Ala Pro Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala
770 775 780

Gly Glu Val Ala Pro Thr Pro Thr Thr Pro Thr Pro Gin Arg Thr Leu
785 790 795 800

Pro Ala

CLAIMS

1. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)
- (b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 121)
- (c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)
- (d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 123)
- (e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 124)
- (f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 125)
- (g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID No. 126)
- (h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 127)
- (i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID No. 128) and
- (j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative

substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

- (a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129) and
 - (b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137), wherein Xaa may be any amino acid.
3. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101 or a complement thereof under moderately stringent conditions.
4. A polypeptide comprising an immunogenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 26-51, 138, 139, 163-183 and 201 or a complement thereof under moderately stringent conditions.
5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.
6. An expression vector comprising a DNA molecule according to claim 5.
7. A host cell transformed with an expression vector according to claim 6.

8. The host cell of claim 7 wherein the host cell is selected from the group consisting of *E. coli*, yeast and mammalian cells.

9. A pharmaceutical composition comprising one or more polypeptides according to any one of claims 1-4 and a physiologically acceptable carrier.

10. A pharmaceutical composition comprising one or more DNA molecules according to claim 5 and a physiologically acceptable carrier.

11. A pharmaceutical composition comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a physiologically acceptable carrier.

12. A vaccine comprising one or more polypeptides according to any one of claims 1-4 and a non-specific immune response enhancer.

13. A vaccine comprising:
a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
a non-specific immune response enhancer.

14. A vaccine comprising:
one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140 and 141, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and
a non-specific immune response enhancer.

15. The vaccine of claims 12-14 wherein the non-specific immune response enhancer is an adjuvant.

16. A vaccine comprising one or more DNA molecules according to claim 5 and a non-specific immune response enhancer.

17. A vaccine comprising one or more DNA sequences recited in SEQ ID Nos.: 3, 11, 12, 140 and 141; and a non-specific immune response enhancer.

18. The vaccine of claims 16 or 17 wherein the non-specific immune response enhancer is an adjuvant.

19. A pharmaceutical composition according to any one of claims 9-11, for use in the manufacture of a medicament for inducing protective immunity in a patient.

20. A vaccine according to any one of claims 12-18, for use in the manufacture of a medicament for inducing protective immunity in a patient.

21. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

22. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6.

23. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO:155).

24. A pharmaceutical composition comprising a fusion protein according to any one of claims 21-23 and a physiologically acceptable carrier.

25. A vaccine comprising a fusion protein according to any one of claims 21-23 and a non-specific immune response enhancer.

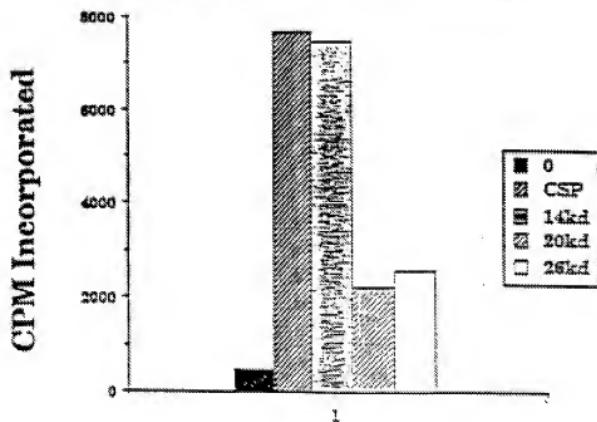
26. The vaccine of claim 25 wherein the non-specific immune response enhancer is an adjuvant.

27. A pharmaceutical composition according to claim 24, for use in the manufacture of a medicament for inducing protective immunity in a patient.
28. A vaccine according to claims 25 or 26, for use in the manufacture of a medicament for inducing protective immunity in a patient.
29. A method for detecting tuberculosis in a patient, comprising:
- contacting dermal cells of a patient with one or more polypeptides according to any one of claims 1-4; and
 - detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.
30. A method for detecting tuberculosis in a patient, comprising:
- contacting dermal cells of a patient with a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
 - detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.
31. A method for detecting tuberculosis in a patient, comprising:
- contacting dermal cells of a patient with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and
 - detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.
32. The method of any one of claims 29-31 wherein the immune response is induration.

33. A diagnostic kit comprising:
 - (a) a polypeptide according to any one of claims 1-4; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
34. A diagnostic kit comprising:
 - (a) a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
35. A diagnostic kit comprising:
 - (a) a polypeptide encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200 and 203; and
 - (b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
36. A diagnostic kit comprising:
 - (a) a fusion protein according to any one of claims 21-23; and
 - (b) apparatus sufficient to contact said fusion protein with the dermal cells of a patient.
37. A fusion protein according to claim 23 comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 153 and 209.

1/11

D7 T Cell Proliferation



D7 IFNg

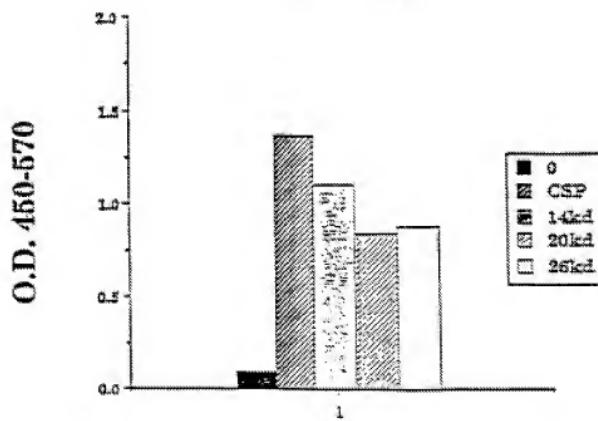
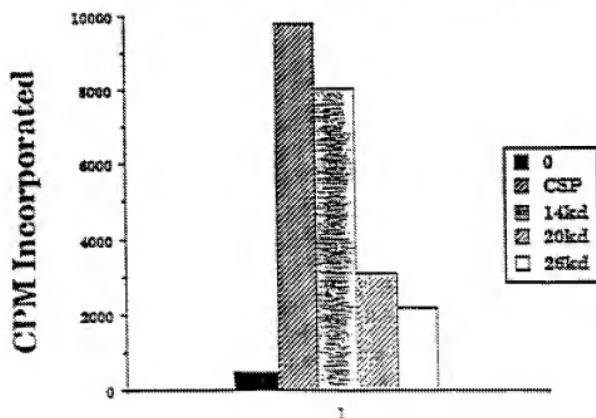


FIG. 1A

D160 T Cell Proliferation



D160 IFNg

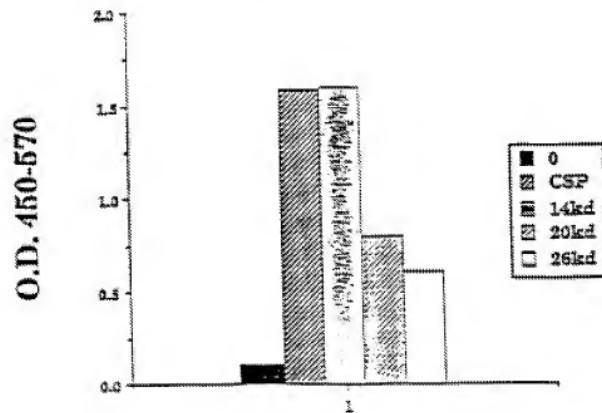
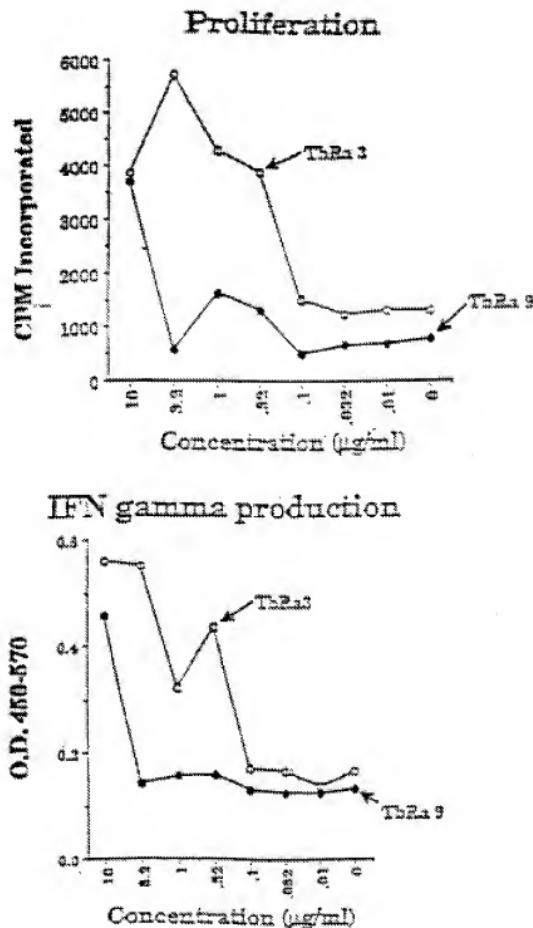
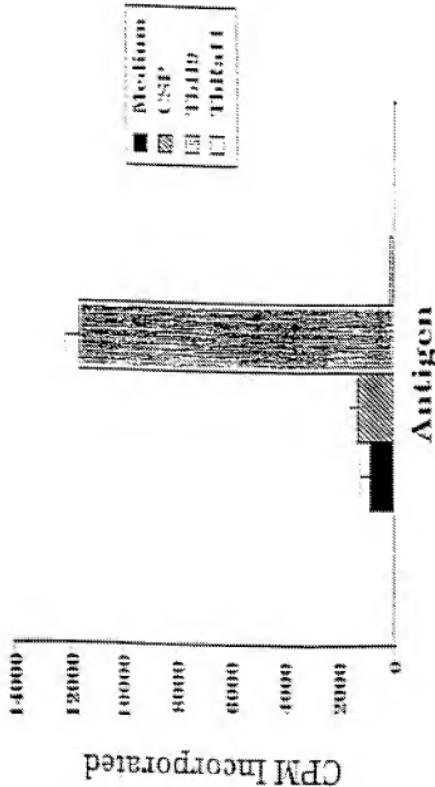


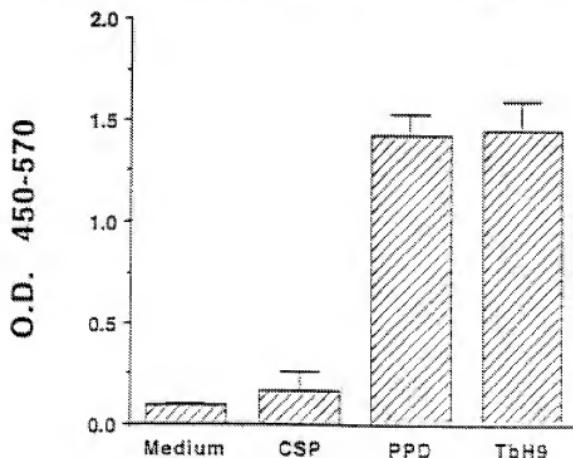
FIG. 1B

**FIG. 2**

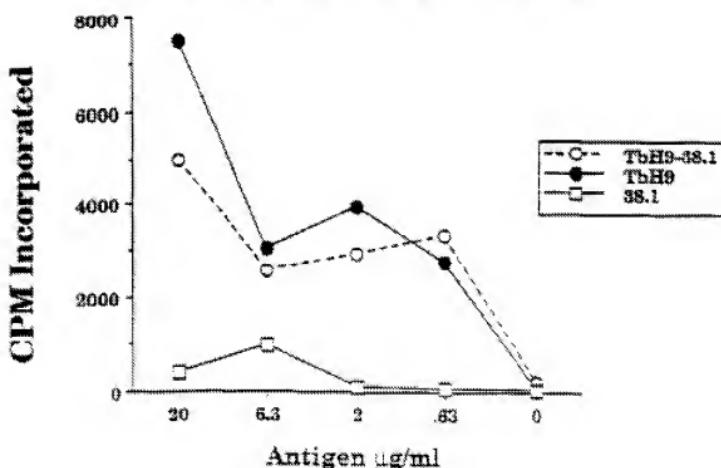
3A*3B**3E**3F**3C**3D**3E**3F**3G**3H**3I**3J**3K**3L**3M**3N**3O**3P**3Q**3R**3S**3T**3U**3V**3W**3X**3Y**3Z**3AA**3AB**3AC**3AD**3AE**3AF**3AG**3AH**3AI**3AJ**3AK**3AL**3AM**3AN**3AO**3AP**3AQ**3AR**3AS**3AT**3AU**3AV**3AW**3AX**3AY**3AZ**3BA**3BB**3BC**3BD**3BE**3BF**3BG**3BH**3BI**3BJ**3BK**3BL**3BM**3BN**3BO**3BP**3BQ**3BR**3BS**3BT**3BU**3BV**3BW**3BX**3BY**3BZ**3CA**3CB**3CC**3CD**3CE**3CF**3CG**3CH**3CI**3CJ**3CK**3CL**3CM**3CN**3CO**3CP**3CQ**3CR**3CS**3CT**3CU**3CV**3CW**3CX**3CY**3CZ**3DA**3DB**3DC**3DD**3DE**3DF**3DG**3DH**3DI**3DJ**3DK**3DL**3DM**3DN**3DO**3DP**3DQ**3DR**3DS**3DT**3DU**3DV**3DW**3DX**3DY**3DZ**3EA**3EB**3EC**3ED**3EE**3EF**3EG**3EH**3EI**3EJ**3EK**3EL**3EM**3EN**3EO**3EP**3EQ**3ER**3ES**3ET**3EU**3EV**3EW**3EX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX**3FY**3FZ**3GA**3GB**3GC**3GD**3GE**3GF**3GG**3GH**3GI**3GJ**3GK**3GL**3GM**3GN**3GO**3GP**3GQ**3GR**3GS**3GT**3GU**3GV**3GW**3GX*

T cell clone 131TbE9 responds poorly to CSP

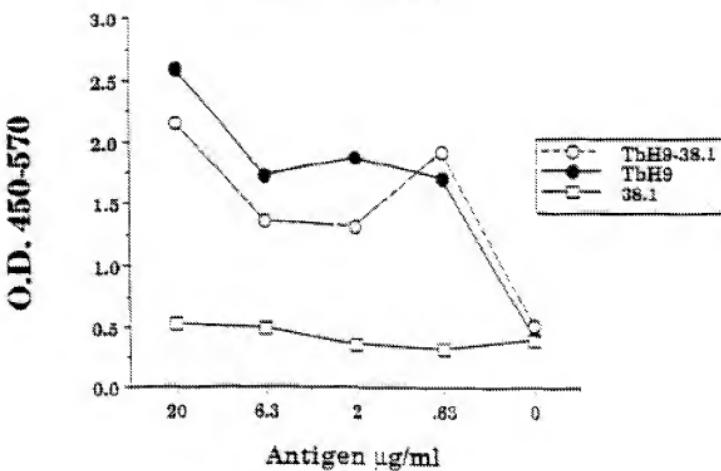


T Cell Clone PPD 800-10 IFNg Production**FIG. 4B**

D131 T Cell Proliferation

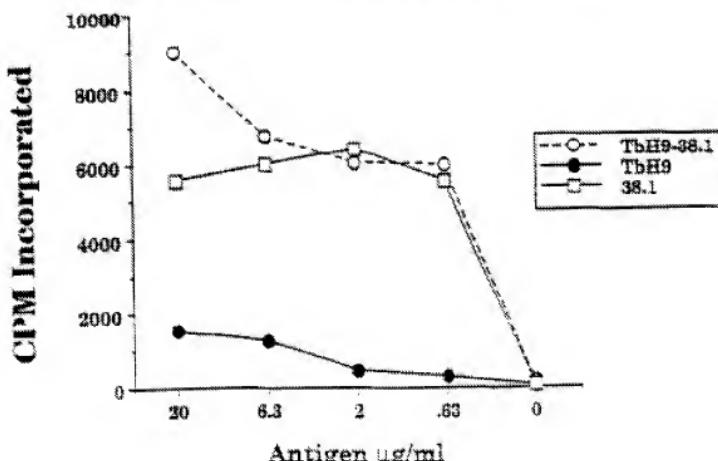


D131 IFNg_g

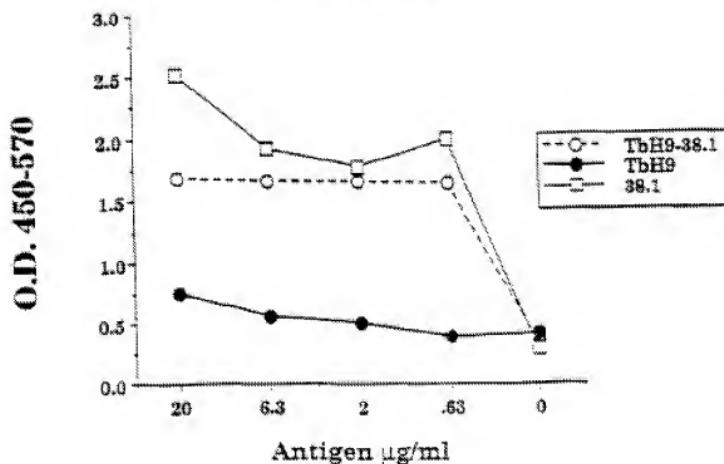


FIGS. 5 A-B

D184 T Cell Proliferation

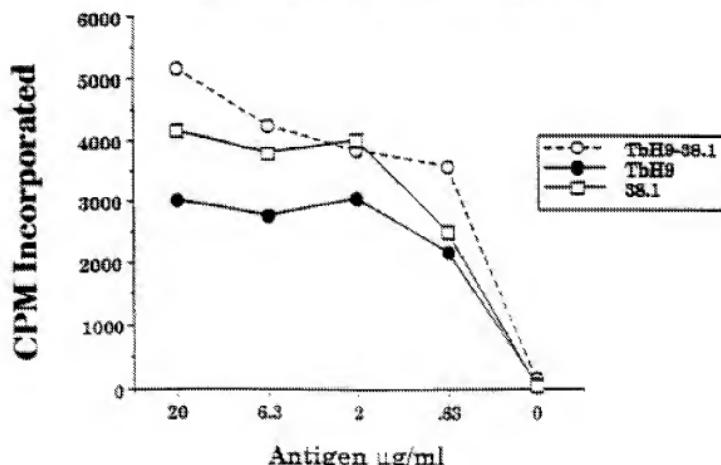


D184 IFNg

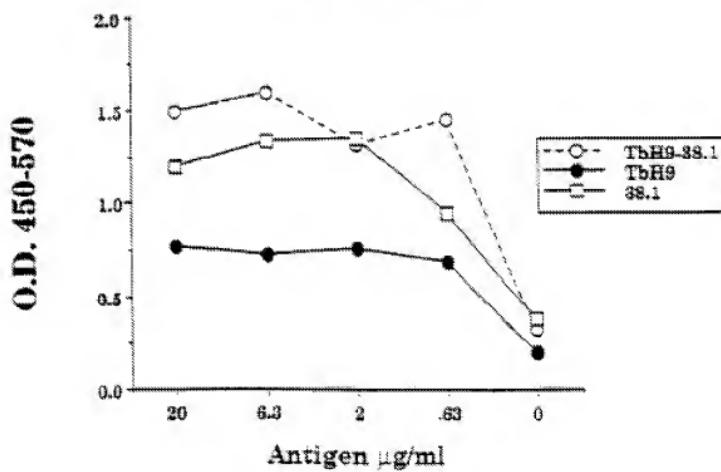


FIGS. 6 A-B

D201 T Cell Proliferation

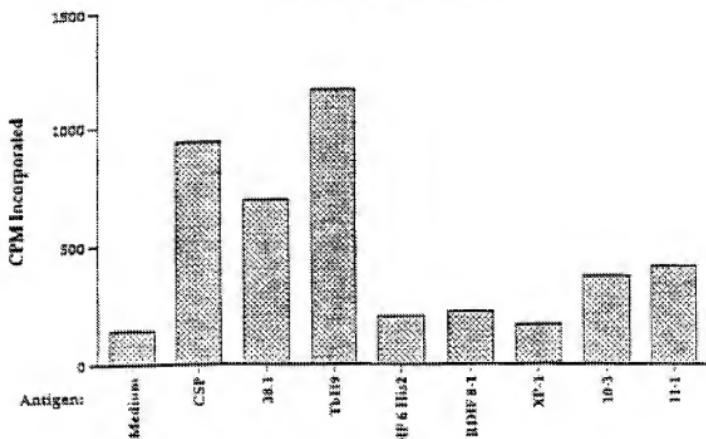


D201 IFNg

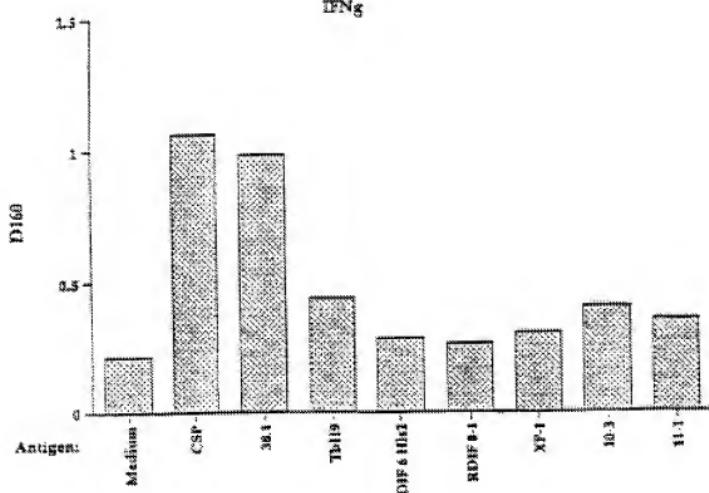


FIGS. 7 A-B

D160 T Cell Proliferation

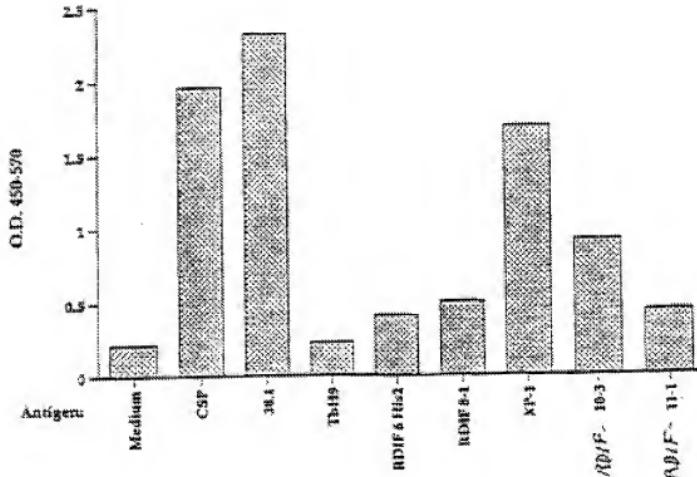
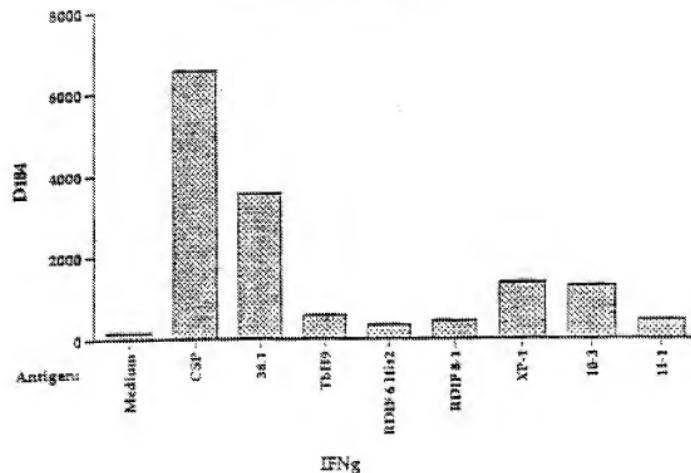


IFNg



FIGS. 8A-B

D184 T Cell Proliferation



FIGS. 9A-B